

U5 – BBDD
NoSQL



QUÉ VAMOS A APRENDER

Consultar información, crear, eliminar y modificar datos utilizando nuevos lenguajes de consulta y nuevos modelos de almacenamiento de datos basados en documentos como XML.

OBJETIVOS DE APRENDIZAJE

- Instalar y utilizar bases de datos embebidas
- Utilizar conectores para acceder a bases de datos
- Establecer conexiones a bases de datos
- Desarrollar aplicaciones para acceder a los datos de la base de datos
- Ejecutar procedimientos de bases de datos

BBDD NoSQL

- No siguen el modelo clásico del sistema de gestión de bases de datos relacionales (RDBMS).
- No usan SQL como lenguaje para sus consultas.
- No se utilizan estructuras fijas para almacenar la información.
- Intentan dar solución a los problemas de las BBDD relacionales:
 - Almacenamiento masivo.
 - Alto desempeño.
 - Procesamiento masivo de transacciones.

NOSQL VS SQL

- Las BBDD relacionales siguen el principio **ACID**:

Atomicity

- Asegurar que la transacción se realice o no, sin quedar a medias ante fallos

Consistency

- Asegurar el estado de validez de los datos en todo momento

Isolation

- Asegurar independencia entre transacciones

Durability

- Asegurar la persistencia de la transacción ante cualquier fallo

- Las BBDD NoSQL se rigen por el principio **BASE**:

Basically Available

- Prioridad de la disponibilidad de los datos

Soft state

- Se prioriza la propagación de datos, delegando el control de inconsistencias a elementos externos

Eventual consistency

- Se asume que inconsistencias temporales progresen a un estado final estable.

VENTAJAS: BBDD NoSQL

- Se ejecutan en máquinas con pocos recursos.
- Escalan horizontalmente para mejorar el rendimiento.
- Permiten manejar gran cantidad de datos.
- No generan cuellos de botella.



DIFERENCIAS BBDDS

NoSQL VS SQL

- No siguen el modelo clásico del sistema de gestión de bases de datos relacionales (RDBMS).
- No usan SQL como lenguaje para sus consultas.
 - Casandra → cql
 - Mongodb → json
 - Bigtable → gql
- No se utilizan estructuras fijas para almacenar la información.
- No suelen permitir operaciones **join**.
- Arquitectura distribuida.
- Intentan dar solución a los problemas de las BBDD relacionales:
 - Almacenamiento masivo.
 - Alto desempeño.
 - Procesamiento masivo de transacciones.

SQL	NoSQL
Es el único lenguaje de consulta de las relaciones	Puede tener diversos lenguajes de consulta (flexible)
Utilizan tablas como modelo de almacenamiento	Utilizan diversos formatos de almacenamientos (JSON, XML, grafos...)
Se usan operaciones JOIN para relacionar/unir datos y tablas	No se suelen hacer operaciones JOIN
Se basan en una arquitectura centralizada	Se basa en una arquitectura distribuida

SQL

Cuando el volumen de mis datos no crece o lo hace poco a poco.

Cuando las necesidades de proceso se pueden asumir en un sólo servidor.

Cuando no tenemos picos de uso del sistema por parte de los usuarios más allá de los previstos.

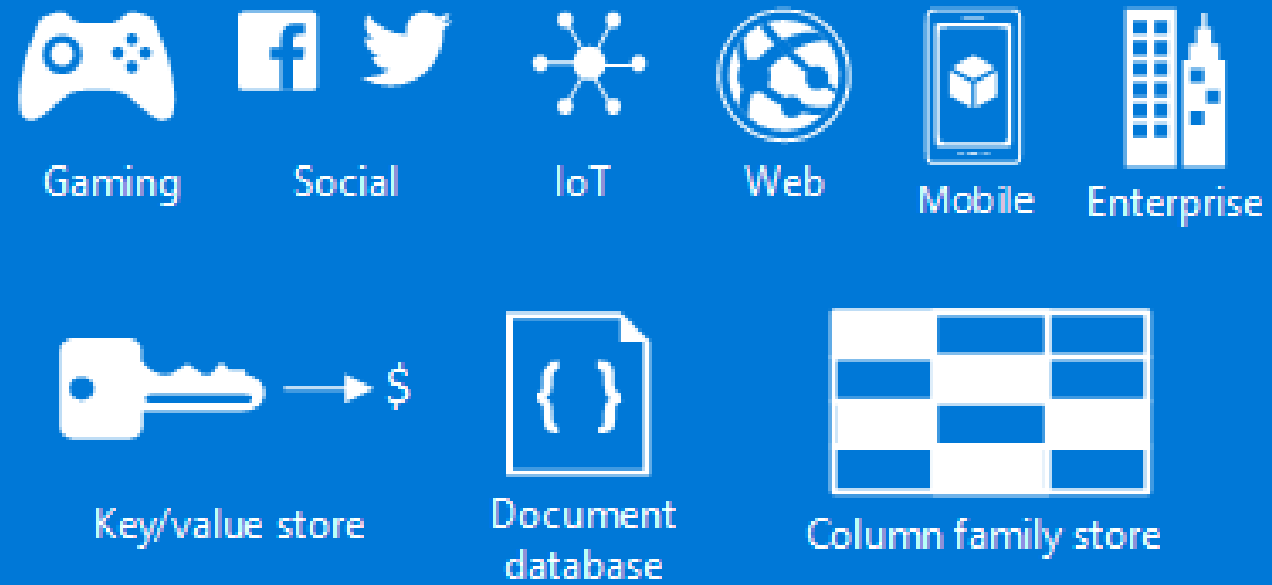
NoSQL

Cuando el volumen de mis datos crece muy rápidamente en momentos puntuales.

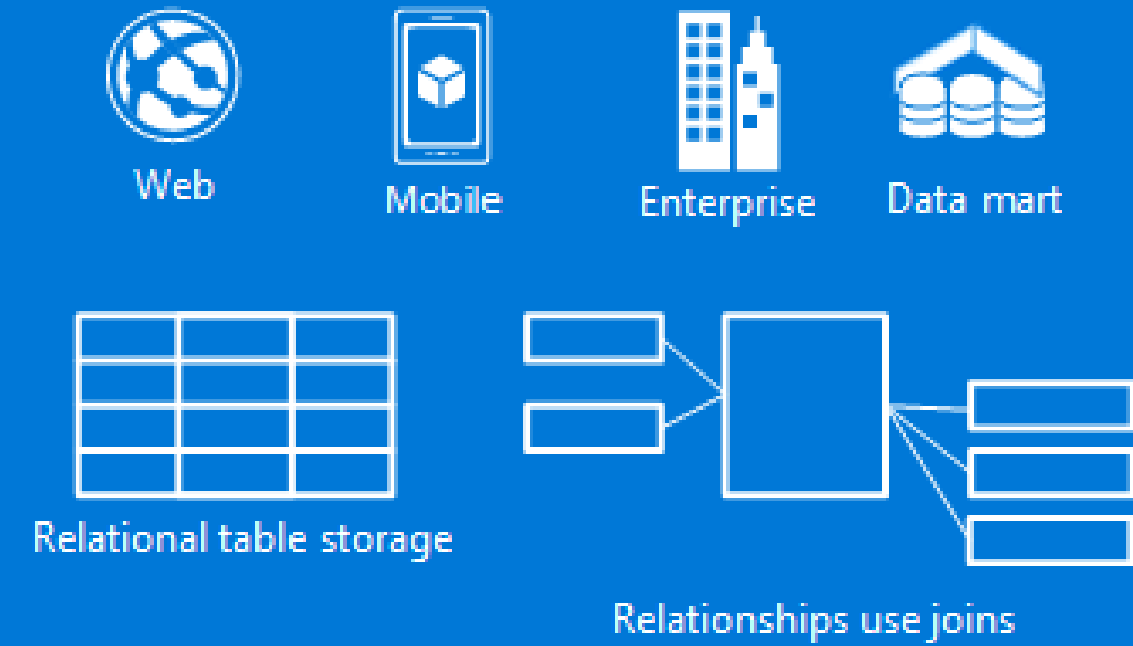
Cuando las necesidades de proceso no se pueden preveer.

Cuando tenemos picos de uso del sistema por parte de los usuarios en múltiples ocasiones.

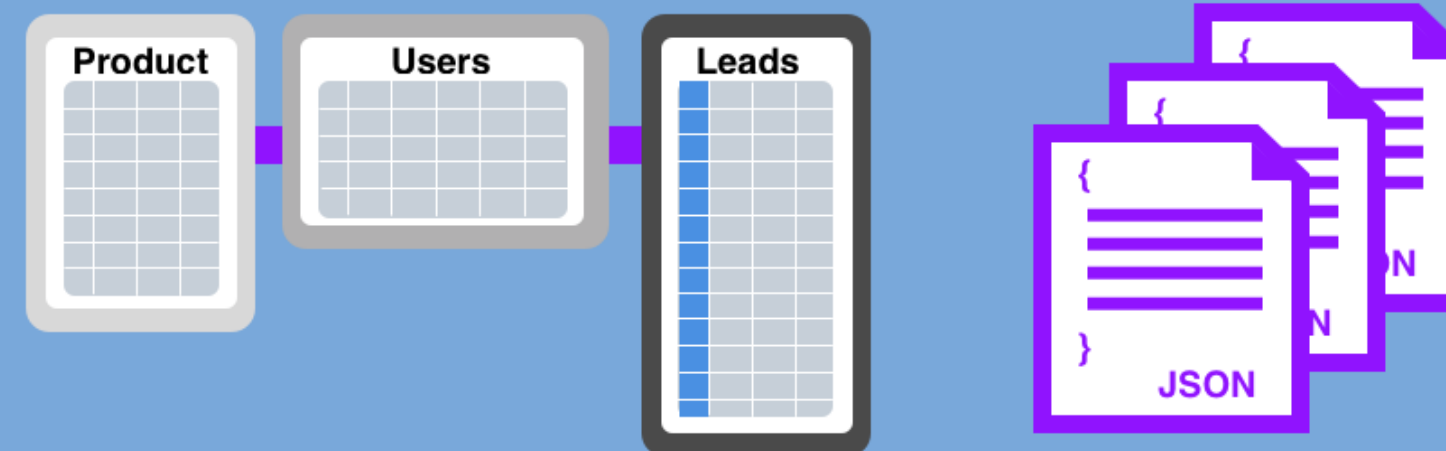
NoSQL



SQL



SQL vs. NoSQL



```
{
  "ID": 1,
  "FIRST": "Frank",
  "LAST": "Weigel",
  "ZIP": "94040",
  "CITY": "MV",
  "STATE": "CA"
}
```

JSON

=

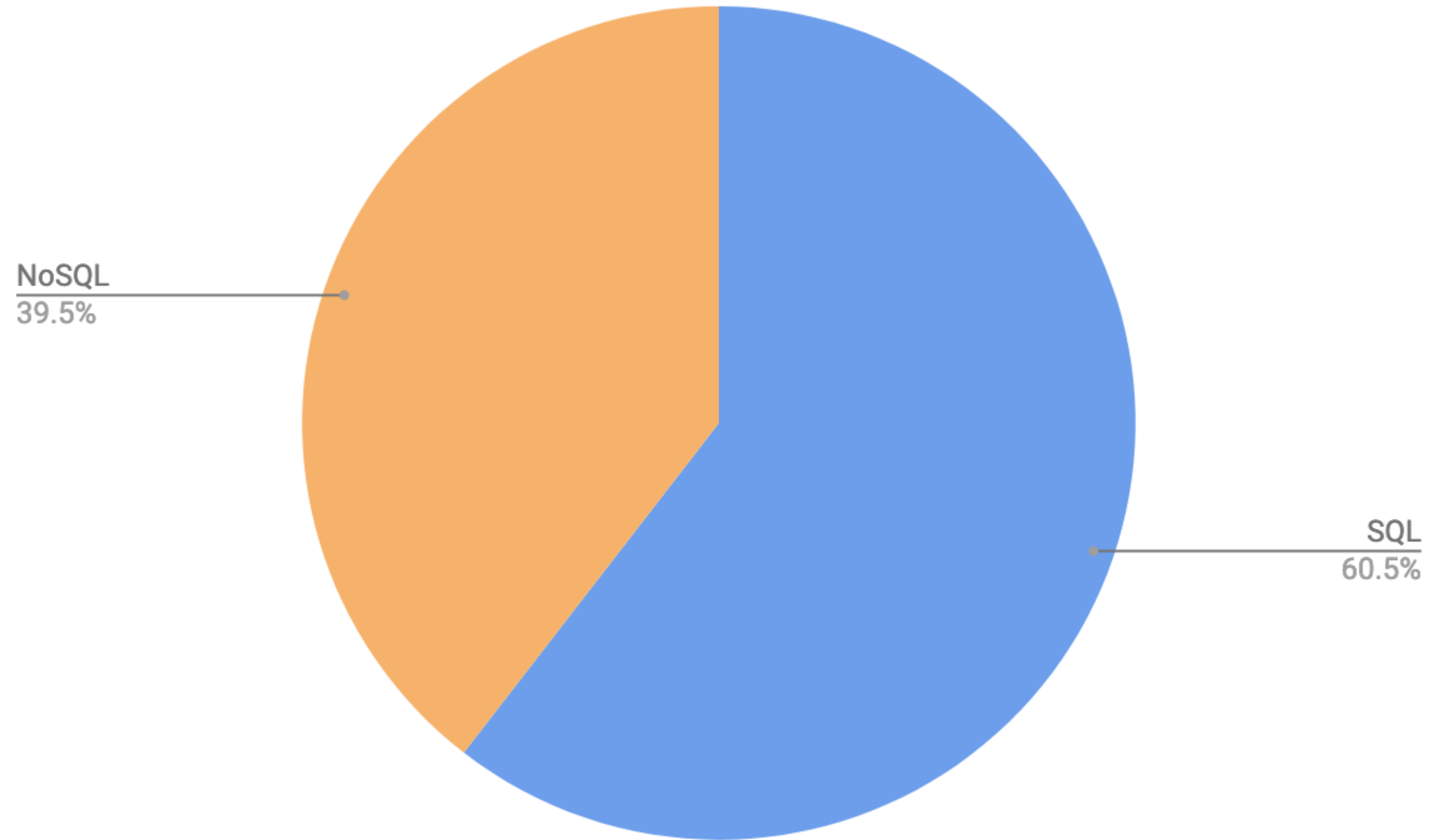
User Info			
KEY	First	Last	ZIP_Id
1	Frank	Weigel	2
2	Ali	Dodson	2
3	Mark	Azad	2
4	Steve	Yen	3

+

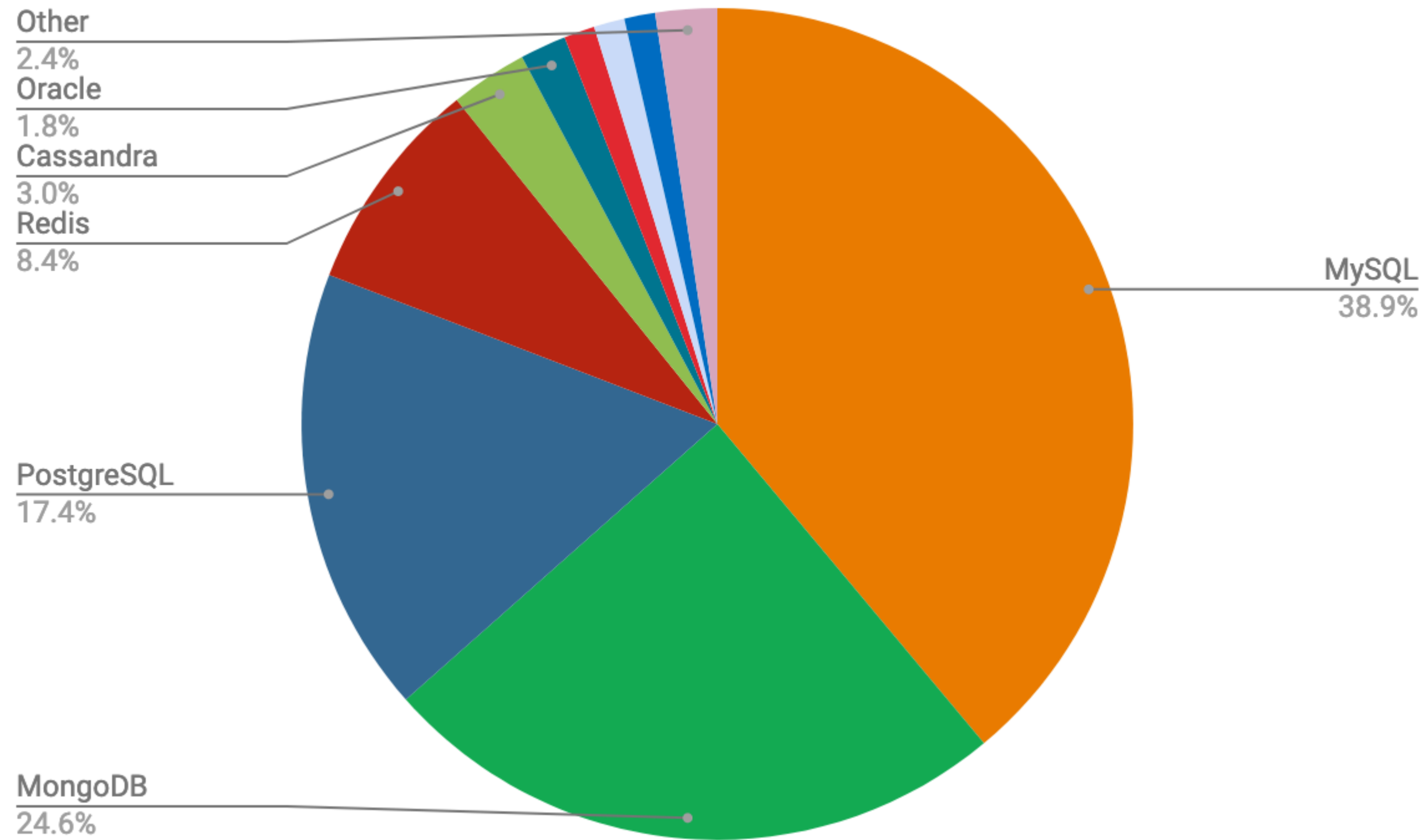
Address Info			
ZIP_Id	CITY	STATE	ZIP
1	DEN	CO	30303
2	MV	CA	94040
3	CHI	IL	60609
4	NY	NY	10010

	row keys	column family "color"	column family "shape"
row	"first"	"red": "#F00" "blue": "#00F" "yellow": "#FF0"	"square": "4"
row	"second"		"triangle": "3" "square": "4"

NoSQL VS SQL



NoSQL VS SQL



BBDD NoSQL

TIPOS

Según el tipo o modelo escogido las BBDD NoSQL se agrupan en cuatro categorías principales:

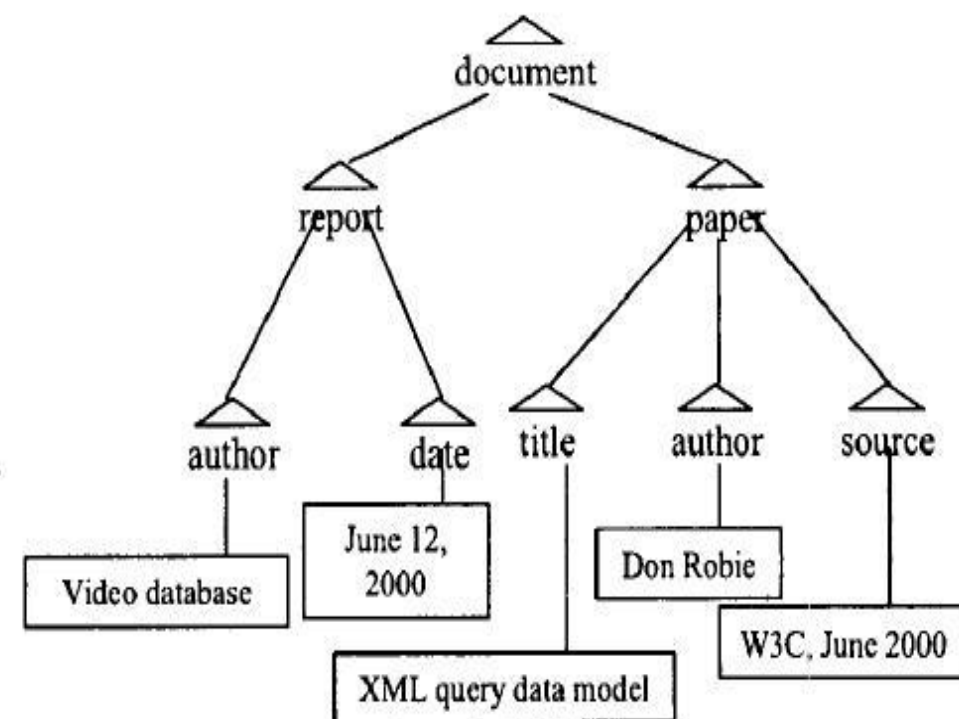
- **Clave/Valor:** Los datos se almacenan y se localizan e identifican usando una clave única y un valor.
- **Columnas:** Como el modelo clave/valor pero la clave se basa en una combinación de columna, fila y marca de tiempo que se utiliza para referenciar conjuntos de columnas.
- **Documentos:** Los datos se almacenan en documentos que encapsulan la información en formato XML, YAML o JSON.
- **Grafos:** Se sigue un modelo de grafos que extiende entre múltiples máquinas. La información se representa como nodos de un grafo y sus relaciones como sus aristas.

BBDD NATIVAS XML

Las BBDD nativas XML no poseen campos, ni almacenan datos, lo que almacenan son documentos XML.

- Define un modelo lógico para un documento XML y almacena y recupera los documentos según este modelo.
- Tiene una relación transparente con el mecanismo de almacenamiento, que debe de incorporar las características ACID de cualquier SGBD.
- Incluye un número arbitrario de niveles de datos y complejidad.

```
<document>
  <report>
    <author>Video database</author>
    <date>June 12, 2000</date>
  </report>
  <paper>
    <title>XML query data model</title>
    <author>Don Robie</author>
    <source>W3C, June 2000</source>
  </paper>
</document>
```



BBDD NATIVAS XML

VENTAJAS

- Ofrecen un acceso y almacenamiento de información en formato XML sin necesidad de incorporar código adicional.
- Suelen incorporar un motor de búsqueda de alto rendimiento.
- Facilidad para añadir documentos XML al repositorio.
- Permite almacenar datos heterogéneos.



BBDD NATIVAS XML

DESVENTAJAS

- Puede resultar difícil indexar documentos para las búsquedas.
- No suelen ofrecer funciones para la agregación, en muchos casos hay que reproducir todo el documento para modificar una línea.
- Se suele almacenar la información en XML como un documento o como un conjunto de nodos, por lo que formar nuevas estructuras sobre la marcha puede resultar complicado y lento.



BBDDS NATIVAS XML

Ejemplo de soporte de datos XML en Oracle:

- Consulta que devuelve las filas de la tabla EMPLE en formato XML:

```
SELECT XMLELEMENT("EMP_ROW",  
    XMLFOREST(EMP_NO, APELLIDO, OFICIO, DIR, FECHA_ALT  
    SALARIO, COMISION, DEPT_NO)) FILA_EN_XML  
FROM EMPLE;
```

- Creación de una tabla que almacena datos del tipo XML:

```
CREATE TABLE TABLE_XML_PRUEBA(COD NUMBER, DATOS XMLTYPE);
```

- Insertar filas en formato XML:

```
INSERT INTO TABLA_XML_PRUEBA VALUES(1,  
    XMLTYPE(' <FILA_EMP><EMP_NO>123</EMP_NO>  
    <APELLIDO>RAMOS MARTÍN</APELLIDO>  
    <OFICIO>PROFESORA</OFICIO>  
    <SALARIO>1500</SALARIO></FILA_EMP>' ) );
```

BBDD NATIVAS XML

Ejemplo de soporte de datos XML en Oracle:

- Extracción de datos de la tabla, se utilizan expresiones XPath:

```
SELECT EXTRACTVALUE (DATOS, '/FILA_EMP/APELLIDO')  
      FROM TABLA_XML_PRUEBA;
```

```
SELECT EXTRACTVALUE (DATOS, '/FILA_EMP/APELLIDO')  
      FROM TABLA_XML_PRUEBA  
 WHERE EXISTSNODE (DATOS, '/FILA_EMP[EMP_NO=123]')  
                = 1;
```



BBDD NATIVAS XML

Ejemplo de soporte de datos XML en MySQL:

- Si disponemos de una BD en MySql llamada “ejemplo” y desearamos obtener los datos de la tabla “departamentos” en formato XML (usuario y contraseña “ejemplo”). Desde la línea de comandos y en la carpeta donde se encuentra instalado MySQL escribimos la siguiente sentencia:

mysql --xml -u ejemplo -p -e "select * from departamentos;" ejemplo

- **--xml:** se utiliza para producir una salida en XML.
- **-u:** para indicar el nombre de usuario.
- **-p:** para que al pulsar intro nos pida la contraseña.
- **-e:** ejecuta el comando y sale de MySQL.

BBDD NATIVAS XML

Ejemplo de soporte de datos XML en MySQL:

- El resultado de la consulta anterior sería:

```
<?xml version="1.0"?>
<resultset statement="select * from departamentos"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <row>
    <field name="dept_no">10</field>
    <field name="dnombre">CONTABILIDAD</field>
    <field name="loc">SEVILLA</field>
  </row>
  <row>
    <field name="dept_no">20</field>
    <field name="dnombre">INVESTIGACIÓN</field>
    <field name="loc">MADRID</field>
  </row>
  <row>
    <field name="dept_no">30</field>
    <field name="dnombre">VENTAS</field>
    <field name="loc">BARCELONA</field>
  </row>
  <row>
    <field name="dept_no">40</field>
    <field name="dnombre">PRODUCCION</field>
    <field name="loc">BILBAO</field>
  </row>
</resultset>
```

BBDD NATIVAS XML

Ejemplo de soporte de datos XML en MySQL:

- Si deseamos redireccionar la salida al archivo ***departamentos.xml*** escribiremos en la misma línea de la consola de comandos:

mysql --xml -u ejemplo -p -e "select * from departamentos;" ejemplo >departamentos.xml





BASE DE DATOS EXIST

- **eXist** es un SGBD libre de código abierto que almacena datos XML de acuerdo a un modelo de datos XML.
- El **motor de base de datos está completamente escrito en Java** y soporta los estándares de consulta **Xpath, Xquery y XSLT**.
- También soporta indexación de documentos y soporte para la actualización de los datos y para multitud de protocolos (SOAP, XML-PRC, WebDav y REST).
- Con el SGBD se consiguen aplicaciones que permiten ejecutar consultas directamente sobre la BD.



BASE DE DATOS EXIST

- Los documentos XML se almacenan en **colecciones**, las cuales pueden estar anidadas.
- No es necesario que los documentos tengan un DTD o un XML Schema asociado (XSD).
- Dentro de una colección se pueden almacenar documentos de cualquier tipo.
- Los archivos más importantes se guardan en la carpeta ***“eXist/webapp/WEB-INF/data”***, y entre ellos están:
 - **dom.dbx**: almacén central de datos, es un fichero paginado donde se almacenan todos los nodos del documento de acuerdo al modelo DOM de W3C.
 - **collections.dbx**: almacena la jerarquía de colecciones y relaciona esta con los documentos que contiene. Se asigna un identificador único a cada documento de la colección que es almacenado también junto al índice.



BASE DE DATOS EXIST

Instalación (I):

eXist puede funcionar de varios modos:

1. Funcionando como un *servidor autónomo* (ofreciendo servicios de llamada remota a procedimientos que funciona sobre Internet como XML-RCP, WebDAV y REST).
2. Insertado *dentro de una aplicación Java*.
3. En un *servidor J2EE*, ofreciendo servicios XML-RPC, SOAP y WebDAV.





BASE DE DATOS EXIST

Instalación (II):

En el sitio <http://exist-db.org/exist/apps/homepage/index.html> podemos descargar la última versión de la BD. Para clase vamos a instalar la versión **eXist 3.0.RC1.jar**, para lo que es necesario tener instalada la versión *Java 8*.

Podemos proceder con la instalación de dos formas:

1. Invocando al fichero jar desde la línea de comandos “*jar -java eXist 3.0.RC1.jar*”.
2. Haciendo doble clic sobre el icono del fichero jar descargado.

Una vez ejecutado seguimos los pasos que indica el asistente, con cuidado de elegir correctamente el directorio para la instalación, así como a la hora de indicar la contraseña del administrador, que en nuestro caso pondremos **admin**.



Actualización: **eXist 5.0.0**



BASE DE DATOS EXIST

Instalación (III):

Una vez instalada podemos arrancar la base de datos desde el acceso directo que se crea en el escritorio, o también desde el menú de la aplicación seleccionando la opción ***eXist-db (XML) Database***, o podemos ejecutar el fichero **startup.jar** que se encuentra en la carpeta de **eXist/bin**.

Al lanzar la BD puede ocurrir que esta no se inicie y nos muestre una pantalla con el error. Esto ocurre cuando al intentar conectar con la BD, el puerto que utiliza eXist está ocupado por otra aplicación. eXist se instala en un servidor web (jetty) y ocupa el puerto **8080**, como la mayoría de servidores web (Apache, Tomcat, Lampp...)



BASE DE DATOS EXIST

Instalación (IV):

Para resolver este problema, cambiamos el puerto. Para ello se edita el fichero de configuración: `%HOME_EXIST%/tools/jetty/etc/jetty.xml`, en la etiqueta `<Call name="addConnector">` cambiaremos el puerto, y en la propiedad ***SystemProperty*** de la etiqueta `<Set name="port">` escribimos un nuevo valor, por ejemplo 8083.

Cambiamos: `<SystemProperty name="jetty.port" default="8080"/>`

Por: `<SystemProperty name="jetty.port" default="8083"/>`



A la hora de ejecutar startup.bat por primera vez, nos pide cambiar el puerto, por lo que (quizás) no sea necesario modificar este archivo



BASE DE DATOS EXIST

Primeros pasos (I):

Teniendo instalado eXist en C:\eXist-db, para poder inicializar la base de datos debemos ejecutar lo siguiente:

- C:\eXist-db\bin\startup.bat
- C:\eXist-db\bin\client.bat

La base de datos una vez arrancada se puede abrir desde el navegador escribiendo:

<http://localhost:8083/exist/>

Para hacer consultas y trabajar con la base de datos lo podemos hacer desde varios sitios:

- Desde el eXide (**Open eXide**): es una de las herramientas para realizar consultas a documentos de la BD, cargar documentos externos a la BD, crear y borrar colecciones, etc. La URL de exide es: <http://localhost:8083/exist/apps/eXide/index.html>, y desde la pestaña *directory* se puede navegar por las carpetas y documentos de la BD. Para realizar operaciones se utilizarán los iconos de la barra de herramientas.



BASE DE DATOS EXIST

Primeros pasos (II):

Para hacer consultas y trabajar con la base de datos lo podemos hacer desde varios sitios:

- Desde el dashboard (**Open dashboard**): el cuadro de instrumentos es el administrador de aplicaciones de la BD. Soporta aplicaciones y plugins. Las aplicaciones proporcionan su propia interfaz gráfica de usuario web, mientras que los plugins se ejecutan dentro del *dashboard*.
 - Aplicaciones: documentación **eXist-bd Documentation**, entorno de consultas **eXide – Xquery IDE**, aplicación de demostración **eXist-bd Demo Apps**.
 - Plugins: gestor de colecciones **Collections**, gestor de backups **Backup**, gestor de usuarios **User Manager**, gestor de paquetes **Package Manager**.

Más información en <http://exist-db.org/exist/apps/doc/dashboard.xml>



BASE DE DATOS EXIST

Primeros pasos (III):

Para hacer consultas y trabajar con la base de datos lo podemos hacer desde varios sitios:

- Desde el cliente java (**Open Java Admin Client**): es la herramienta que utilizaremos para hacer las consultas. El cliente nos pedirá conexión, con el usuario y la contraseña (admin/admin), y hay que asegurarse de poner correctamente el puerto de la URL ***xmldb:exist://localhost:8083/exist/xmlrpc***.



Podemos acceder al cliente Java a través de la **ruta de instalación\bin\client.bat**

- admin/admin
- Puerto: 8083

<https://exist-db.org/exist/apps/doc/java-admin-client>



BASE DE DATOS EXIST

El cliente de administración de eXist (I):

- Una vez conectado se muestra la ventana del *Cliente de Administración eXist*. **Desde aquí se pueden hacer todo tipo de operaciones sobre la BD** (crear y borrar colecciones, añadir y eliminar documentos a las colecciones, modificar los documentos, crear copias de seguridad, administrar usuarios y realizar consultas XPath...)
- También podemos **navegar por las colecciones** y elegir un contexto a partir de donde se ejecutarán las consultas. También se puede hacer doble clic sobre un documento para modificarlo.
- Si pulsamos el botón de los prismáticos (**Consultar la BD usando XPath**), aparece la ventana de consultas ***Diálogo de consulta***, desde aquí se puede elegir también, el contexto sobre el que ejecutaremos las consultas, así como guardar las consultas y sus resultados en ficheros externos.



BASE DE DATOS EXIST

Cliente de Administración eXist

Fichero Herramientas Conexión Opciones Ayuda

Recurso	Fecha	Propietario	Grupo	Permisos
apps	2019-10-21 17:1...	SYSTEM	dba	crwxr-xr-x
system	2019-10-21 01:5...	SYSTEM	dba	crwxr-xr-x

type help or ? for help.
exist:/db> |

Cliente de Administración de eXist conectado - admin@xmldb:exist://localhost:8083/exist/xmlrpc

Query: podemos ejecutar, guardar consultas y sus resultados en ficheros externos

Administrar usuarios



BASE DE DATOS EXIST

El cliente de administración de eXist (II):

Para subir una colección a la BD lo primero que haremos es conectarnos con admin.

En la ventana del *Cliente* pulsamos el botón *Almacena uno o más ficheros en la base de datos*, y en la ventana que aparece seleccionamos la carpeta a subir y pulsamos el botón *Seleccionar los ficheros o directorios*.

Una vez pulsado aparece una ventana de transferencia de datos, y al cerrarla, la colección con los documentos se creará en la BD, y lo hará dentro de la colección donde se llamó a la subida de documentos (en principio se subirán dentro del contexto ***exist:/db>***).

Para hacer consultas sobre esta colección hacemos doble clic sobre ella para seleccionarla, y una vez dentro pulsamos el botón *Consultar la base de datos usando XPath*.



BASE DE DATOS EXIST

Cliente de Administración eXist

Fichero Herramientas Conexión Opciones Ayuda

Almacena uno o más ficheros en la base de datos

Recurso	Fecha	Propietario	Grupo	Permisos
apps		SYSTEM	dba	crwxr-xr-x
system		SYSTEM	dba	crwxr-xr-x

type help or ? for help.
exist:/db>

Cliente de Administración de eXist conectado - admin@xmldb:exist://localhost:8083/exist/xmlrpc

Seleccionar los ficheros o directorios a almacenar

Buscar en: ColeccionesXML

Elementos re...

Escritorio

Documentos

Este equipo

Red

ColeccionPruebas

Nombre de archivo: JNIDAD5-BBDDs NoSQL\RecursosUnidad5\ColeccionesXML

Archivos de tipo: Todos los Archivos

Seleccionar los ficheros o directorios a almacenar

Cancelar



BASE DE DATOS EXIST

Cliente de Administración eXist

Fichero Herramientas Conexión Opciones Ayuda

Recurso	Fecha	Propietario	Grupo	Permisos
ColeccionPruebas		admin	dba	crwxr-xr-x
apps		SYSTEM	dba	crwxr-xr-x
system		SYSTEM	dba	crwxr-xr-x



Cliente de Administración eXist

Fichero Herramientas Conexión Opciones Ayuda

Recurso	Fecha	Propietario	Grupo	Permisos
NUEVOS_DEP.xml		admin	dba	-rw-r--r--
departamentos.xml		admin	dba	-rw-r--r--
departamentosn...		admin	dba	-rw-r--r--
empleados.xml		admin	dba	-rw-r--r--
productos.xml		admin	dba	-rw-r--r--
sucursales.xml		admin	dba	-rw-r--r--
universidad.xml		admin	dba	-rw-r--r--
zonas.xml		admin	dba	-rw-r--r--



Cliente de Administración eXist

Fichero Herramientas Conexión Opciones Ayuda

Recurso	Fecha	Propietario	Grupo	Permisos
NUEVOS_DEP.xml		admin	dba	-rw-r--r--
departamentos.xml		admin	dba	-rw-r--r--
departamentosn...		admin	dba	-rw-r--r--
empleados.xml		admin	dba	-rw-r--r--
productos.xml		admin	dba	-rw-r--r--
sucursales.xml		admin	dba	-rw-r--r--
universidad.xml		admin	dba	-rw-r--r--
zonas.xml		admin	dba	-rw-r--r--

type help or ? for help.
 exist:/db> cd "ColeccionPruebas"
 exist:/db/ColeccionPruebas> cd ..
 exist:/db> cd "ColeccionPruebas"
 exist:/db/ColeccionPruebas>

Cliente de Administración de eXist conectado - admin@xmldb:exist://localhost:8083/exist/xmlrpc
 exist:/db> cd "ColeccionPruebas"
 exist:/db/ColeccionPruebas> cd ..
 exist:/db> cd "ColeccionPruebas"
 exist:/db/ColeccionPruebas>

Cliente de Administración de eXist conectado - admin@xmldb:exist://localhost:8083/exist/xmlrpc

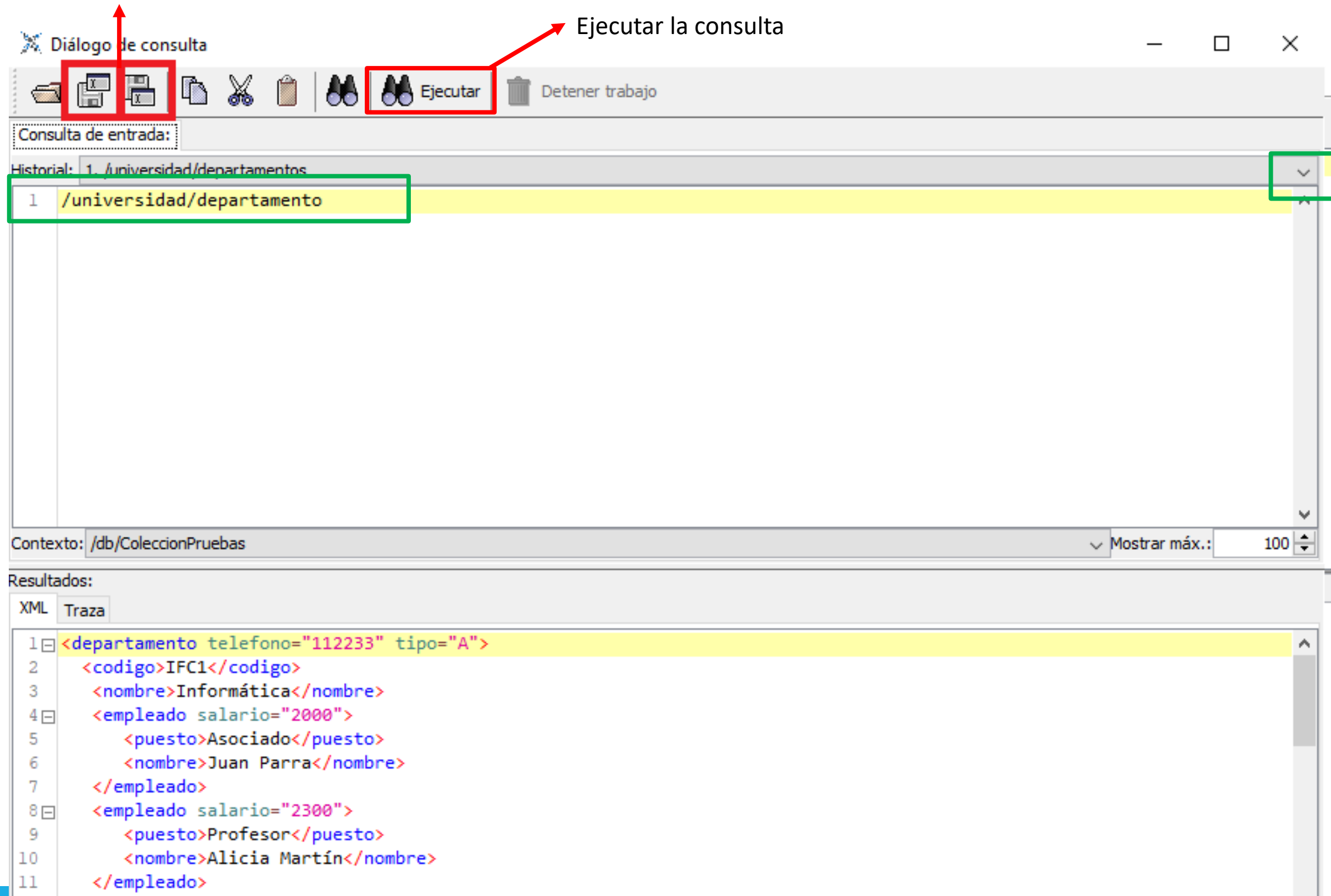
universidad.xml

```

<universidad>
  <departamento telefono="112233" tipo="A">
    <codigo>IFC1</codigo>
    <nombre>Informática</nombre>
    <empleado salario="2000">
      <puesto>Asociado</puesto>
      <nombre>Juan Parra</nombre>
    </empleado>
    <empleado salario="2300">
      <puesto>Profesor</puesto>
      <nombre>Alicia Martín</nombre>
    </empleado>
  </departamento>
  <departamento telefono="990033" tipo="A">
    <codigo>MAT1</codigo>
    <nombre>Matemáticas</nombre>
    <empleado salario="1000">
      <puesto>Asociado</puesto>
      <nombre>Juan Parra</nombre>
    </empleado>
  </departamento>
</universidad>
  
```

universidad.xml desde xmldb:exist://localhost:8083/exist/xmlrpc Line: 1 Column: 1

1. Guardar la consulta
2. Guardar resultados de consultas





BASE DE DATOS EXIST

Actividad

Una vez configurado el cliente, realiza las siguientes consultas con la colección *"ColeccionPruebas"*:

- Dentro del recurso "universidad.xml" obtener el departamento y todos los demás datos.
- Dentro del recurso "universidad.xml" obtener sólo el código de cada departamento.
- Dentro del recurso "universidad.xml" obtener sólo el nombre de los empleados

/universidad/departamento/

/universidad/departamento/código

/universidad/departamento/empleado/nombre



BASE DE DATOS EXIST

Actividad

Crea un archivo XML que guarde los siguientes datos:

CICLO	ASIGNATURA	DATOS		
		MODALIDAD	NOC_DIU	NUM_ALU
DAW	ED	A	DIU	22
SMR	RL	A	DIU	24
	MME	B	DIU	24
DAM	ED	A	NOC	24
	AD	A	NOC	10
ADFI	OPI	A	NOC	20

Añádelo a la colección “*ColeccionPruebas*”.

Crea tres consultas diferentes para acceder a los datos: (1)modalidad, (2)nocturno o diurno y (3)número de alumnos.