**Taller de base de datos**

**Alumno:**

Cansino Rosas Erick

**Unidad 1**

**Fecha:**

22/11/2019

**Docente:**

Nava Castro Pioquinto Hugo

Un **Sistema Gestor de Base de Datos** (**SGBD**) o DataBase Managenent System (**DBMS**) es un sistema que permite la **creación, gestión y administración de bases de datos**, así como la **elección y manejo de las estructuras necesarias** para el almacenamiento y búsqueda de información del modo más eficiente posible. Se clasifican según la forma en que administran los datos: Relacionales y no relacionales.

Relacionales

|  |  |  |  |
| --- | --- | --- | --- |
| SGBD | LICENCIA | VENTAJAS/DESVENTAJAS | DESCRIPCIÓN |
| MySQL | GNU GPL o comercial | \*Facilidad de uso y gran rendimiento  \*Facilidad para instalar y configurar  \*Soporte multiplataforma  \*Soporte SSL  -----------------------------------------  \*Escalabilidad, no trabaja de manera eficiente con BD que superan un determinado tamaño | Es un SGBD multihilo y multiusuario utilizado en la gran parte de las páginas web actuales. Además es el más usado en aplicaciones creadas como software libre. |
| MariaDB | GNU GPL | \*Aumento de motores de almacenamiento  \*Gran escalabilidad  \*Seguridad y rapidez en transacciones  \*Extensiones y nuevas características relacionadas con su aplicación para Bases de datos NoSQL.  -----------------------------------------  \*Pequeñas incompatibilidades en la migración de MariaDB y MySQL.  \*Pequeños atrasos en la liberación de versiones estables. | Totalmente compatible con MySQL |
| SQLite | Open Source | \*Permite transacciones sin necesidad de un servidor ni configuraciones  \*El tamaño, al tratarse de una biblioteca, es mucho menor que cualquier SGBD  \*Reúne los cuatro criterios ACID (Atomicidad, Consistencia, Aislamiento y Durabilidad) logrando gran estabilidad  \*Gran portabilidad y rendimiento  -----------------------------------------  \*Escalabilidad, no soporta BD muy grandes. | Es una biblioteca escrita en c que implementa un SGBD |
| PostgreSQL | BSD | \*Control de Concurrencias Multiversión (MVCC)  \*Flexibilidad en cuanto a lenguajes de programación  \*Multiplataforma  \*Dispone de una herramienta (PGAdmin) muy fácil e intuitiva para la administración de las bases de datos.  \*Robustez, Eficiencia y Estabilidad.  -----------------------------------------  \*Lentitud para la administración de bases de datos pequeñas ya que está optimizado para gestionar grandes volúmenes de datos. | Este sistema gestor de base de datos relacional está orientado a objetos. |
| Microsoft SQL Server | Código cerrado  Microsoft EULA | \*Soporte exclusivo por parte de Microsoft.  \*Escalabilidad, estabilidad y seguridad.  \*Posibilidad de cancelar consultas.  \*Potente entorno gráfico de administración que permite utilizar comandos DDL y DML.  \*Aunque es nativo para  -------------------------------------------  \*El precio (Estándar, Developer, Enterprise, SQL Azure) | Es un sistema gestor de bases de datos relacionales basado en el lenguaje Transact-SQL, capaz de poner a disposición de muchos usuarios grandes cantidades de datos de manera simultánea. |
| Oracle | privativa | Soporte de transacciones.  Estabilidad.  Escalabilidad.  Multiplataforma.  -------------------------------------------  Es el coste del software (SE, SE1, SE2, PE, LE, Enterprise Edition) | Considerado el SGBD por excelencia para el mundo empresarial, por su robustez y completitud. |

Una base de datos no relacional (NoSQL) es aquella base de datos que:

* No requiere de estructuras de datos fijas como tablas
* No garantiza completamente las características ACID (características de los parámetros que permiten clasificar las transacciones de los SGBD: Atomicidad, Consistencia, Aislamiento y Durabilidad)
* Escala muy bien horizontalmente.
* Se utilizan en entornos distribuidos que han de estar siempre disponibles y operativos y que gestionan un importante volumen de datos.

No relacionales:

|  |  |  |  |
| --- | --- | --- | --- |
| SGBD | LICENCIA | VENTAJAS/DESVENTAJAS | DESCRIPCIÓN |
| MongoDB | Open Source | Indexación y replicación  Balanceo de carga  Almacenamiento en ficheros  Consultas ad hoc  Escalabilidad horizontal  --------------------------------------------  Como desventaja principal, MongoDB no es un SGBD adecuado para realizar transacciones complejas. | Es un SBGD NoSQL orientado a ficheros que almacena la información en estructuras BSON con un esquema dinámico que permite su facilidad de integración. |
| Redis | BSD | \*El principal uso de este SGBD es para el almacenamiento en memoria caché y la administración de sesiones.  \*Atomicidad y persistencia  \*Gran velocidad  \*Simplicidad  \*Multiplataforma | Está basado en el almacenamiento clave-valor. Podríamos verlo como un vector enorme que almacena todo tipo de datos, desde cadenas, hashses, listas, etc. |
| Cassandra | Apache Licencia 2.0 | \*Multiplataforma  \*Propio lenguaje de consultas (CQL)  \*Escalado lineal y horizontal  \*Es un SGBD distribuido  \*Utiliza una arquitectura peer-to-peer | Es un SGBD NoSQL distribuido y masivamente escalable. |

<https://revistadigital.inesem.es/informatica-y-tics/los-gestores-de-bases-de-datos-mas-usados/>

<https://www.diarlu.com/gestores-bases-datos/>

Cantidad de usuarios en BD:

la cantidad de usuarios que accede a la información del BD, se configura a través de parámetros dirigidos hacia las consultas, actualizaciones y conexiones, pero todo depende en gran medida del servidor disponible para el SGBD, Esta estimación puede ayudarle a determinar la configuración de hardware que necesitará para realizar lo siguiente:

* Conseguir el rendimiento que necesitan las aplicaciones.
* Asegurar la cantidad física adecuada de espacio en disco necesario para almacenar los datos y los índices.

Todo gestor SGBD permite ejecutar las siguientes operaciones:

Definir la estructura de datos: en la definición de los datos se guarda una descripción con metadatos de la estructura de datos en el diccionario del sistema. se utiliza para definir los datos el Data Definition Language (DDL).

Definir derechos: todos los lenguajes de bases de datos proporcionan una sintaxis que permite otorgar o retirar permisos. En este se utiliza el Data Control Language (DCL) o lenguaje de control de datos.

Definir condiciones de integridad: por condiciones de integridad se entienden los requisitos de estado que se exigen a un banco de datos. Si se definen condiciones para su integridad, el BD garantiza que se cumplan en todo momento. Se habla entonces de un estado consistente. Una condición básica de integridad en una base de datos relacional es, por ejemplo, que cada registro (tuplas) pueda identificarse de forma inequívoca.

Definir transacciones: cuando se lleva a un BD de un estado consistente a otro diferente se habla de transacción. Estas transacciones contienen una serie de instrucciones que deben ejecutarse siempre de forma íntegra. Si una se interrumpe, el BD vuelve a su estado original (Rollback). Cada transacción comienza con una orden para crear una conexión con el BD a la que siguen otras que inician las operaciones de datos en sí, así como un paso de comprobación (Commit) que asegura la integridad del BD. Las operaciones que pongan en peligro la integridad de la tabla no se consignan (committed), es decir, no se escriben en la base de datos de forma permanente. Por último, se cierra la conexión con el BD. Al vocabulario del lenguaje de bases de datos con el que se manipulan los datos se le conoce como Data Manipulation Language (DML).

Definir vistas: las llamadas views son vistas virtuales de un subconjunto de los datos de una tabla. Para crear una vista, el SGBD genera una tabla virtual (relación lógica) sobre la base de las tablas físicas. En estas vistas pueden emplearse las mismas operaciones que se utilizarían en tablas físicas. Las más habituales son aquellas que filtran determinadas filas (consulta de selección) o columnas (vista de columnas) de una tabla, así como las que conectan diversas tablas entre sí (vista de conjunto).

Funciones de un servidor:

Mediante el uso de un cliente de base de datos, se puede acceder a la información que se guarda en las diferentes bases de datos. Una vez el cliente ha accedido mediante un usuario, contraseña y nombre de host, se le permiten realizar diferentes tareas, dependiendo del nivel de privilegios que posea.

¿Qué server de BD debo elegir?

Existen muchas variables a tomar en cuenta cuando vamos a elegir un motor o servidor de base de datos, entre las cuales podemos destacar:

* Tamaño de las bases de datos
* Cantidad de registros por tabla
* Porcentaje de operaciones de escritura y lectura
* Sistema operativo del servidor
* Simultaneidad de usuarios
* Alta disponibilidad y balanceo
* Asegura que el sistema operativo corra adecuadamente a la memoria RAM requerida.
* Si ves que ejerce mucha presión sobre el CPU o ves un alto rendimiento de el mismo, ten en cuenta que puede ser riesgoso o correrá más lento de lo común.
* Observa detalladamente el rendimiento y consumo sobre tu ordenador del servidor de datos, de tal manera que estés al tanto de todo lo que hay que saber sobre tu servidor y sepas manejarlo correctamente a la hora de implementarlo en tu web.
* Mantén vigilado el almacenamiento del servidor de datos, para que el almacenamiento sea el suficiente tanto para hoy como para los días próximos, de forma que no se llene rápidamente.

Servidores web con bases de datos: más sencilla de arquitectura utiliza un solo servidor dedicado para alojar su servidor de base de datos y un servidor web, scripts de sitios web y otros archivos. Aunque esta opción es la más sencilla, puede considerarse como una opción de alto rendimiento siempre que el funcionamiento de su web y base de datos no sea demasiado intenso.

Clúster de servidores: Es una solución de alojamiento de alta disponibilidad que permite que servidores específicos ejecuten servicios específicos. El clúster consta de una parte del hardware para servidores (o más) que aloja sus bases de datos, sin necesidad de efectuar otros servicios como ejecutar un servidor web. Eso lo maneja otra parte del hardware para servidores.

Storage Attached Network (SAN): Red de área de almacenamiento hace que un conjunto de dispositivos de almacenamiento estén disponibles en múltiples servidores de bases de datos, de tal manera que los dispositivos aparecen como dispositivos conectados localmente al sistema operativo.

<https://www.ionos.mx/digitalguide/hosting/cuestiones-tecnicas/bases-de-datos-relacionales/>

<https://blog.infranetworking.com/servidor-base-de-datos/>

<https://www.youngeek.com/servidor-base-datos/>

### MySQL

Sistema de gestión de **base de datos relacional multiusuario y multihilo**. Desarrollada inicialmente por Sun Microsystems, actualmente su propietario es Oracle Corporation. Software libre y tipo de licencia GNU GPL para aquellos proyectos que se adapten a este tipo de licencia (permite usar, compartir y modificar el software), con la opción de poder adquirir una de carácter más privativo previo pago (parte del código está bajo copyright). Página Oficial: <http://www.mysql.com>(

Características más destacadas:

* Escrito en C y C++.
* Sistemas de almacenamiento transaccional y no transaccional.
* Tablas hash en memoria como tablas temporales.
* Servidor disponible como software externo para un entorno cliente/servidor.
* Usa GNU Automake, Autoconf, y Libtool para portabilidad.
* Registros de longitud fija y longitud variable.
* Diversos tipos de columnas: FLOAT, DOUBLE, CHAR, VARCHAR, TEXT, BLOB, DATE, TIME, DATETIME, TIMESTAMP, YEAR, SET y ENUM.
* Soporta numerosos Sistemas Operativos
* Reserva de memoria basada en threads.

### PostgreSQL

Es un sistema de gestión de **bases de datos objeto-relacional**, distribuido bajo licencia BSD y con disponibilidad completa sobre el código fuente. Utiliza el modelo cliente/servidor, y usa multiprocesos en lugar de multihilos, para garantizar que el fallo en un proceso no afecte al resto de procesos en ejecución. Página Oficial: <http://www.postgresql.org/>

Características principales:

* Alta concurrencia con el modelo MVCC (Acceso Concurrente Multiversión) que permite el acceso a una tabla en la que se están realizando tareas de escrituras sin bloquearla para el resto de procesos.
* Base de Datos ACID.
* Funciones y procedimientos almacenados en varios lenguajes de programación, entre los que se encuentran PL/pgSQL, PL/Perl, PL/Python y PL/Tcl.
* Acceso encriptado vía SSL.
* Integridad referencial (es aquella que garantiza que un registro o fila existente se relaciona con otras entidades válidas existentes en la base de datos).
* Soporta copias de seguridad en caliente.
* Límites de tamaño de la estructura de Base de Datos:
  + Base de Datos: Sin límite
  + Tabla: 32 TB
  + Fila: 1,6 TB.
  + Campo: 1 GB
  + Filas por Tabla: Ilimitado
  + Nº Columnas por Tabla: Entre 250 y 1600 dependiendo del tipo de dato.
  + Nº de índices por Tabla: Ilimitado.

<https://academiaandroid.com/conexion-base-datos-externa-android-sgbd/>

Oracle

Costo de capacitación

Uno de los costos más grandes de implementación es el de la capacitación de los usuarios, para que estos operen y mantengan un entorno Oracle.

Oracle recomienda que se tenga cuando menos un administrador de base de datos. La capacitación consiste en cursos de niveles I y II, con un costo aproximado de 120,000 pesos

Costo de instalación

El proceso de una implementación Oracle for WorkGroups, es extenso. Adicionalmente a la instalación inicial, de una base de datos, respaldos, carga inicial de los datos y configuración de procedimientos con un costo aproximado de 57,500 pesos

Beneficio

Se puede ejecutar en cualquier plataforma

Soporta todas las funciones que se esperan de un servidor

Permite el uso de particiones para mejorar el rendimiento

El software del servidor se puede ejecutar en multitud de sistemas operativos

Oracle es un BD con orientación hacia internet

Requerimientos de hardware y software

RAM 512 MB

Memoria virtual 1 GB

Espacio en disco duro 1.5 GB

Sistema operativo: Windows 7, Windows server, Linux, unix

Arquitectura del SO: 32/64 bits

Protocolo de red TCP/IP y TCP/IP con SSL

|  | Caracteres\* por campo de serie de tipos de datos | | Bytes de datos por fila de base de datos (registro) | | Número máximo de columnas por tabla de base de datos† |
| --- | --- | --- | --- | --- | --- |
|  | Oracle 11g | 4000\* bytes por columna | | Ningún límite definido | 1000 |

<https://www.ibm.com/support/knowledgecenter/es/SSNW2F_5.2.1/com.ibm.p8.pe.user.doc/bpfcg008.htm>

MySQL

costo

En este caso el software es libre para todo aquel que quiera implementar un BD, pero si se desea adquirirlo para una empresa y que desempeñe una tarea en específico se debe adquirir una licencia especial la cual permitirá desarrollar la actividad deseada.

Beneficio

Puede ser ejecutado en una maquina con escasos recursos sin ningún problema.

Facilidad de configuración e instalación

Entonces soporta gran variedad de sistemas operativos

Baja probabilidad de corromper datos incluso si no se dañan dentro del gestor sino en el sistema en el que están.

Su conectividad, velocidad y seguridad hace de MySQL altamente apropiado para acceder a BD en internet.

Requerimientos de hardware y software al mínimo

RAM 512 MB

Memoria virtual 1 GB

Espacio en disco duro 1 GB

Tamaño BD “sin límite”

SO: Windows server, Windows 7, Linux, unix

Arquitectura del SO: 32/64 bits

Protocolo de red TCP/IP y TCP/IP con SSL

En las instancias de base de datos MySQL en Amazon RDS, el límite máximo de almacenamiento aprovisionado restringe el tamaño de una tabla a un tamaño máximo de 16 TB cuando se usan espacios de tabla file-per-table de InnoDB. Este límite también restringe el espacio de tabla del sistema a un tamaño máximo de 16 TB. El espacio de tabla file-per-table de InnoDB (en los que las tablas están cada una en su propio espacio de tabla) se habilitan de manera predeterminada para las instancias de base de datos MySQL en Amazon RDS.

<https://docs.aws.amazon.com/es_es/AmazonRDS/latest/UserGuide/MySQL.KnownIssuesAndLimitations.html>

PostgreSQL

COSTOS Y SOPORTE

 El costo de licencia y de soporte con dominantes. Basados en precios publicados, nuestro análisis es que PostgreSQL es 12.5 veces más barato que el sistema de gestión de base de datos relacionales líder en el mercado. Dado un clúster de 6 servidores geográficamente redundantes, conectados usando replicación en tiempo real, con un maestro y 5 esclavos de solo lectura, hemos calculado el precio por licencia y suporte en:

2ndQuadrant - $60,000

Comercial vendedor - $750,000

Note que, a diferencia de 2ndQuadrant, el vendedor comercial hacer una recarga por cada socket, por lo que cada que los requerimientos de rendimiento del sistema aumentan con el tiempo, también aumenta la disparidad en el coste total de propiedad.

VERSIONES

PostgreSQL empezó con la versión 6, para mantener consistencia con el control de versiones de Berkeley, como una señal de reconocimiento a la importante contribución hecha por el equipo de Stonebraker.

Los aproximadamente 4 años y 5 correspondientes versiones principales (de la 7.0 a la 7.4) que marcaron la era 7. \* trajeron más mejoras.

PLATAFORMAS SOPORTADAS POR POSTGRESQL.

En general, cualquier plataforma moderna compatible con Unix, como Linux, Mac OS X y FreeBSD, puede ejecutar PostgreSQL, así como Microsoft Windows. PostgreSQL no depende en el comportamiento de ninguna arquitectura en particular, por lo que casi se puede asegurar que cualquier arquitectura soportada por estos sistemas operativos es soportada por PostgreSQL.

 LICENCIA DE POSTGRESQL

PostgreSQL es distribuido bajo una licencia similar a BSD y MIT. Les permite a los usuarios hacer cualquier cosa que quieran con el código, incluyendo la reventa de los binarios sin el código. La única restricción es que no puede hacer legalmente responsable a los desarrolladores originales por los problemas con el software (aunque esto es, por su puesto, aparte de cualquier acuerdo que pueda hacer con 2ndQuadrant).

Las licencias para el uso en producción no cuestan, las licencias para desarrolladores no cuestan, y las licencias para empresas asociadas no cuestan. Si necesita expandir su uso, puede hacerlo sin volver a planear el presupuesto por lo que su planificación no tiene que ser tan exacta.

PostgreSQL es fácil de instalar por lo que no hay un costo adicional escondido por instalación.

En resumen, este es el factor principal que influye en el coste total de propiedad, pero de ninguna manera es el único.

<https://pardodaniels.wordpress.com/2014/08/21/25/>

<http://perezlili.blogspot.com/2014/08/base-de-datos-postgresql.html>

MANUAL DE INSTALACIÓN

DE POSTGRESQL

# Introducción

El Manual de Instalación de PostgreSQL en su versión 9.3, contiene las especificaciones más relevantes para la instalación de estos softwares. PostgreSQL es un Sistema de gestión de bases de datos relacional orientado a objetos y de distribución libre, por lo que nos será útil para la materia de Taller de Bases de Datos.

La finalidad de este instructivo fundamentalmente es la de proporcionar al lector la lógica con la que se ha desarrolla este proceso de instalación, debido al similar mecanismo de instalación que se lleva a cabo en las diferentes versiones, por tal razón se considera importante documentar esta información, así mismo mostrar las ventajas y recursos que ofrece un software libre.

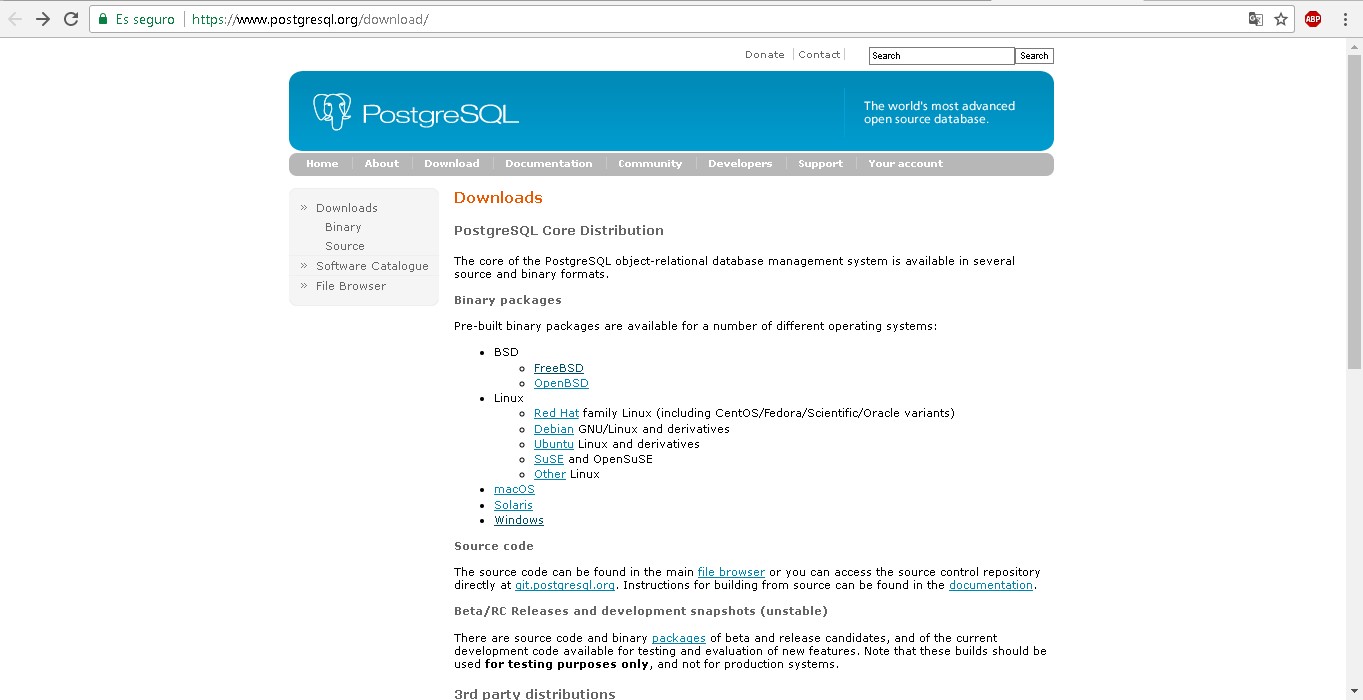
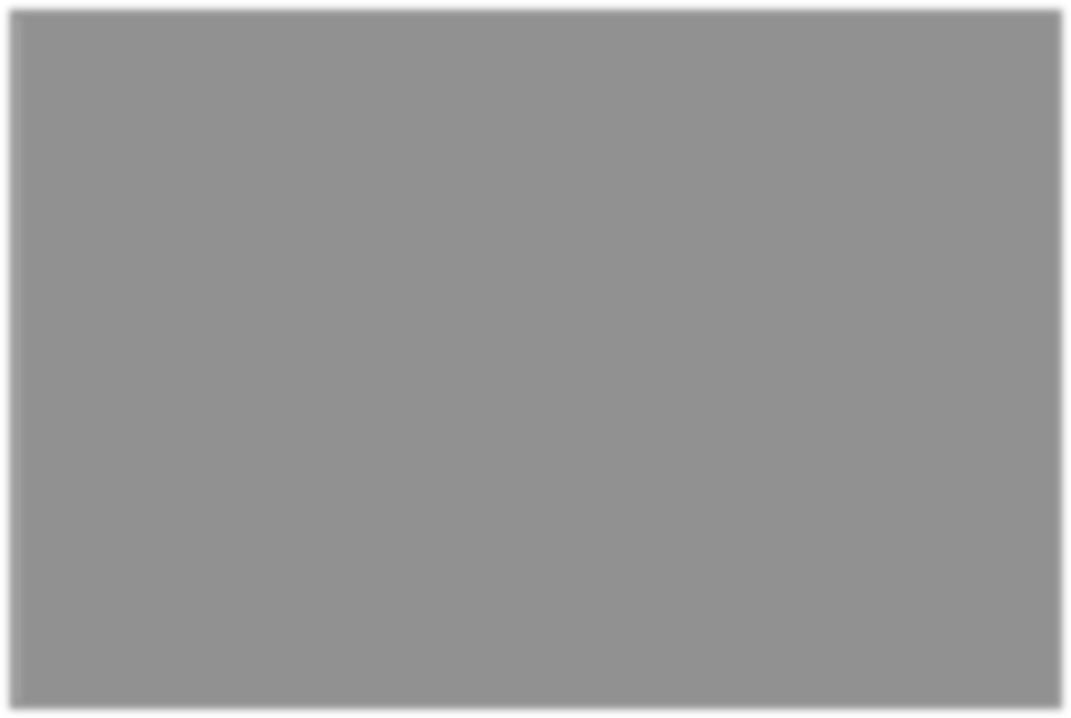
# Objetivos

* + Proporcionar una guía que permita al usuario que acceda a este manual mayor facilidad al momento de realizar la Instalación de PostgreSQL.

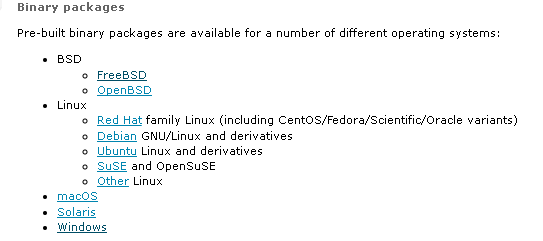
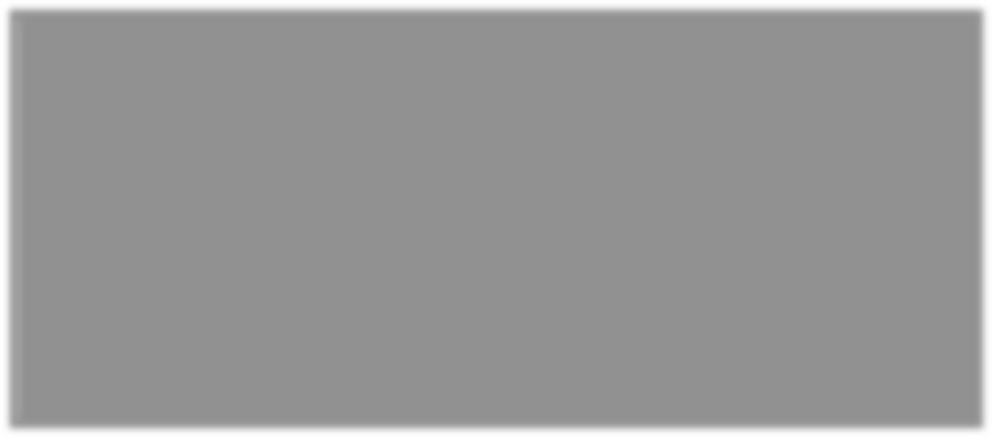
# Contenido

* 1. Instalación de PostgreSQL

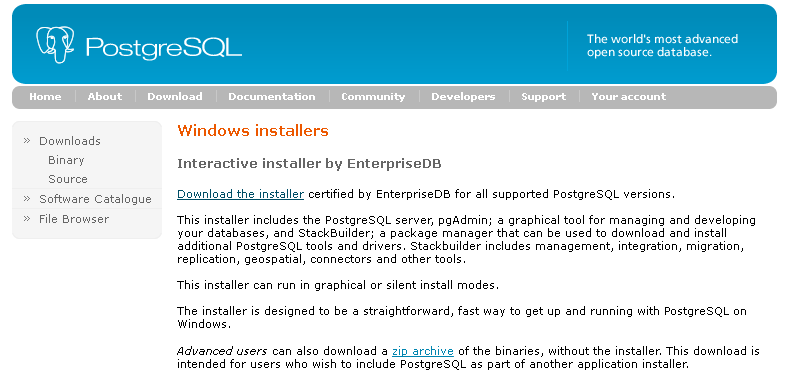
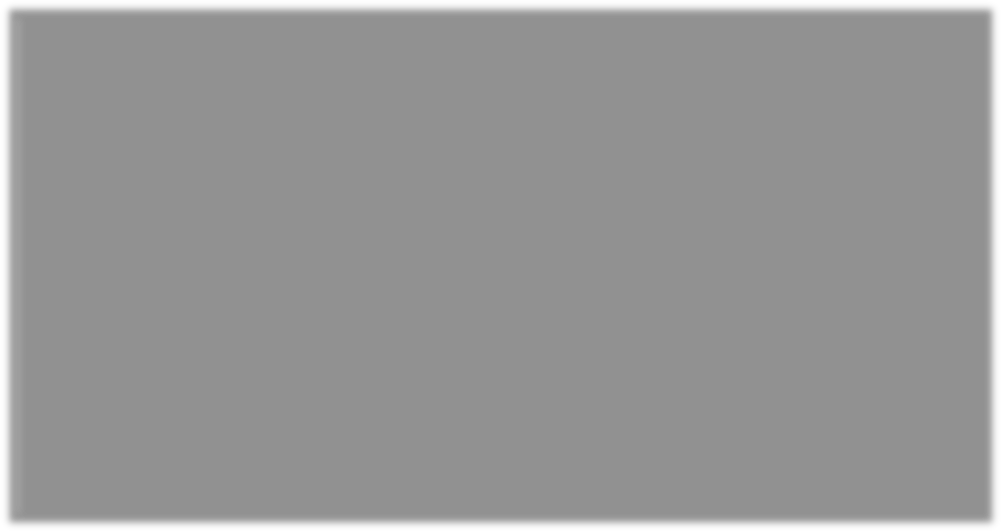
Para obtener el instalador de PostgreSQL accedemos a su página de distribución <https://www.postgresql.org/download/>.



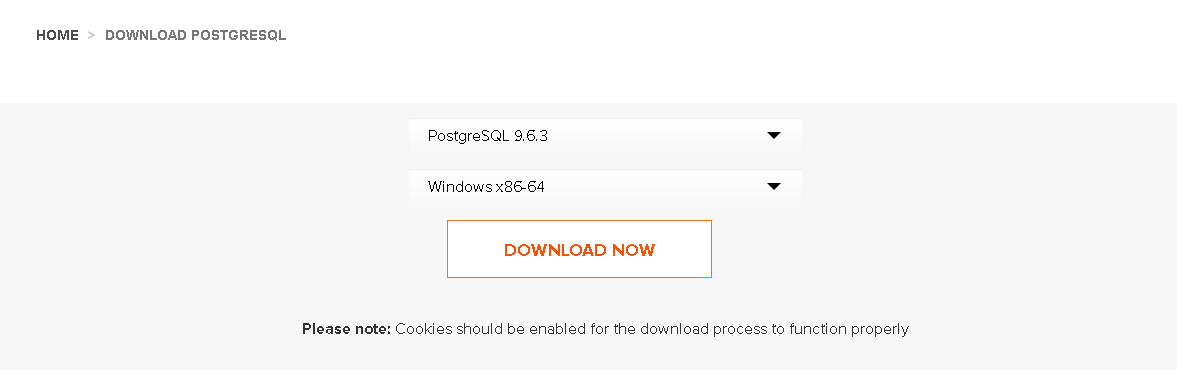
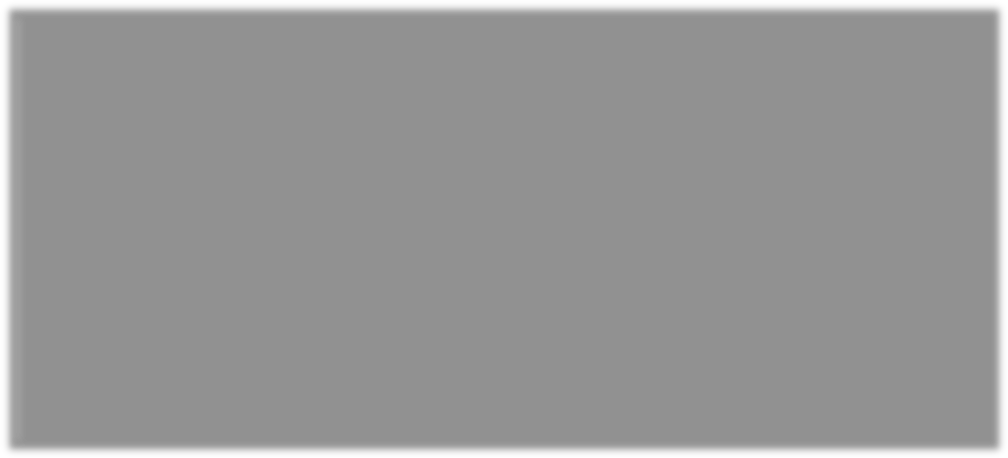
Escogemos la versión que se adapte nuestro sistema operativo, en este tutorial se selecciona la versión compatible con Windows.



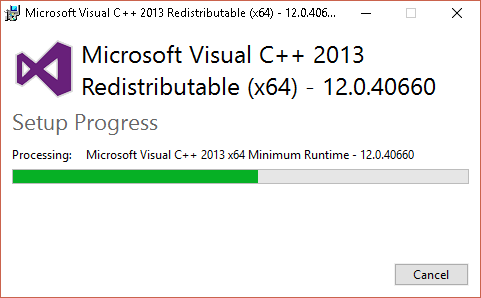
Seleccionamos Download the installer



En nuestro caso escogemos la versión 9.3 de PostgreSQL para Windows de 64 bits (dependiendo de nuestra máquina) y presionamos DOWNLOAD NOW.



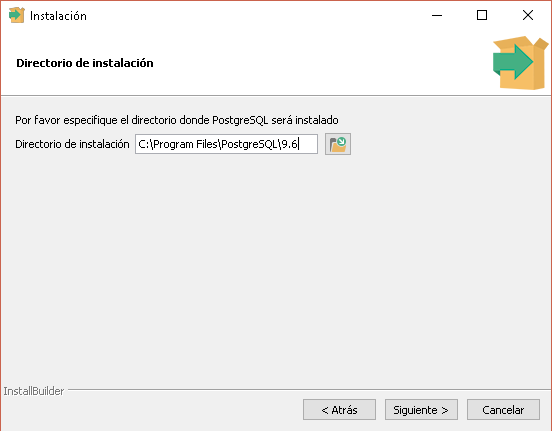
Al ejecutar el instalador se inicia el proceso de instalación de PostgreSQL se empieza a instalar visual C++ debido a que PostgreSQL está escrito en C++.



Aparece la pantalla de Bienvenida para la instalación de PostgreSQL, presionamos el botón “Siguiente” para continuar con el proceso de instalación.

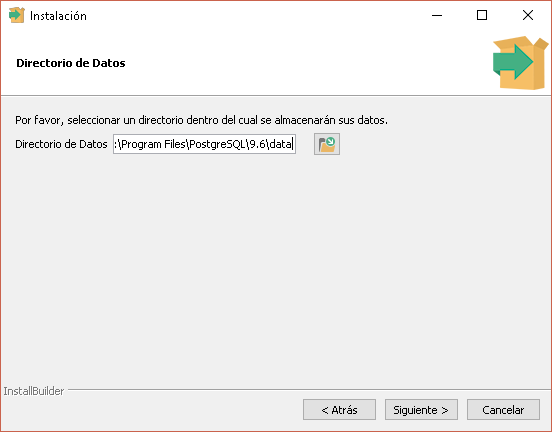


El instalador solicita que se especifique el directorio donde se almacenaran los archivos para esta instalación se deja la opción por defecto, damos click en “Siguiente”.



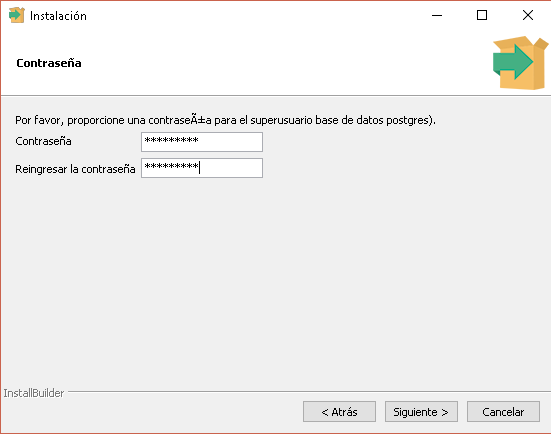
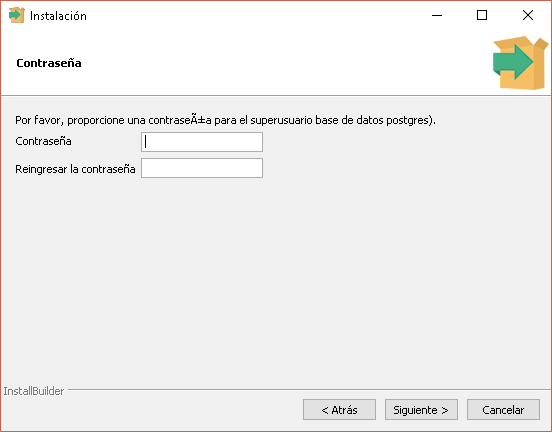
3

De igual forma el instalador pide especificar un directorio para el almacenamiento de Datos, y dejamos la opción por defecto proporcionada por el instalador, presionamos “Siguiente”.

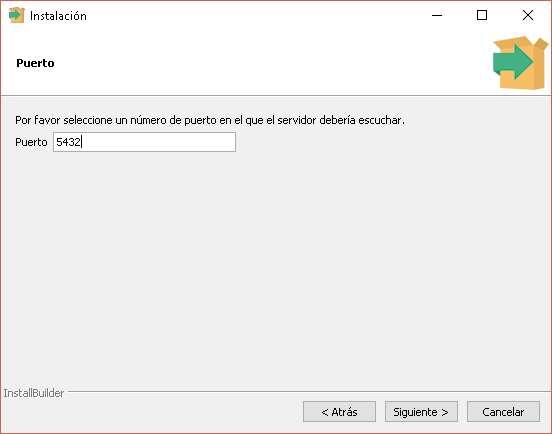


3

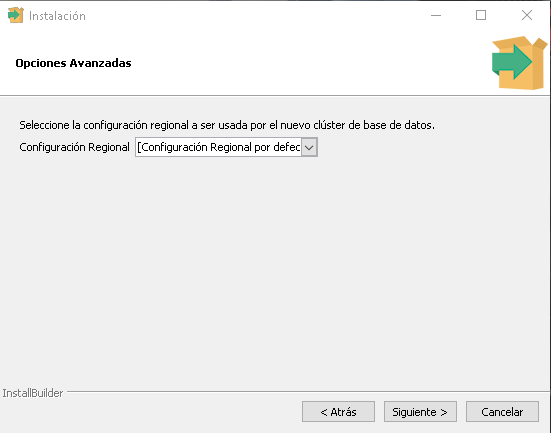
Se nos pedirá proporcionar una contraseña de súper usuario, una vez establecida presionamos “Siguiente”.



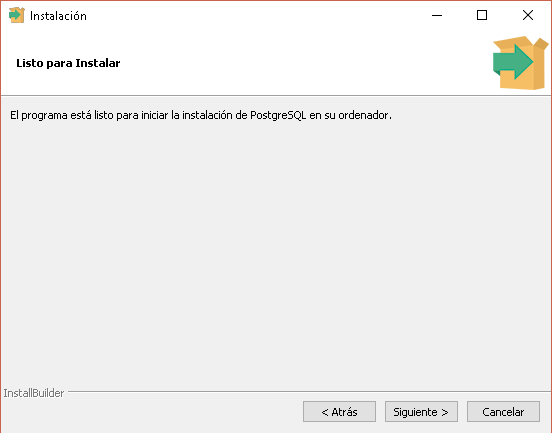
Se proporciona un número de puerto en el que el servidor debería escuchar, se deja el puerto por defecto 5432 y presionamos “Siguiente”.



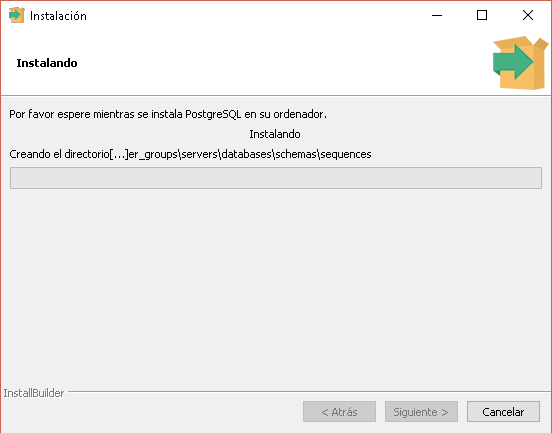
Se debe seleccionar la configuración regional usada por el nuevo clúster de base de datos, se deja la que viene por defecto, presionamos “Siguiente”.



Una vez establecidos los parámetros el programa de instalación está listo para la instalación de PostgreSQL en el computador, damos click en “Siguiente”.



El proceso de instalación está cargando los diferentes componentes, esto puede tardar unos minutos.



Al terminar el proceso de instalación de PostgreSQL, Se selecciona casilla Stack Builder el cual es usado para descargar e instalar herramientas adicionales, controladoras y aplicaciones para complementar la instalación de PostgreSQL, y presionamos “Terminar”

Modelo/esquma ER simple de una preparatoria.

un alumno tiene materias, una materia tiene alumnos; se relacionan por carga académica

una materia tiene profesores, un profesor tiene materias; se relacionan por grupos

con esto se determina que:

un grupo tiene muchas materias, una materia tiene muchos grupos

un grupo tiene muchos profesores, un profesor tiene muchos grupos

un grupo tiene muchos salones

Al estar interrelacionados se determina que un profesor tiene muchos alumnos y que muchos alumnos tienen un profesor. Con este diagrama ER, se determina que la base de datos posee demasiada inestabilidada debido a que todas sus relaciones son de muchos a muchos, aplicando normalización se descompondría en más tablas. La alternativa sería añadir otra capa de tablas puente. Materias\_grupo, grupo\_salones, grupos\_profesores.

Sistema basado en referencias, si quieres la info accedes a través de relaciones.