

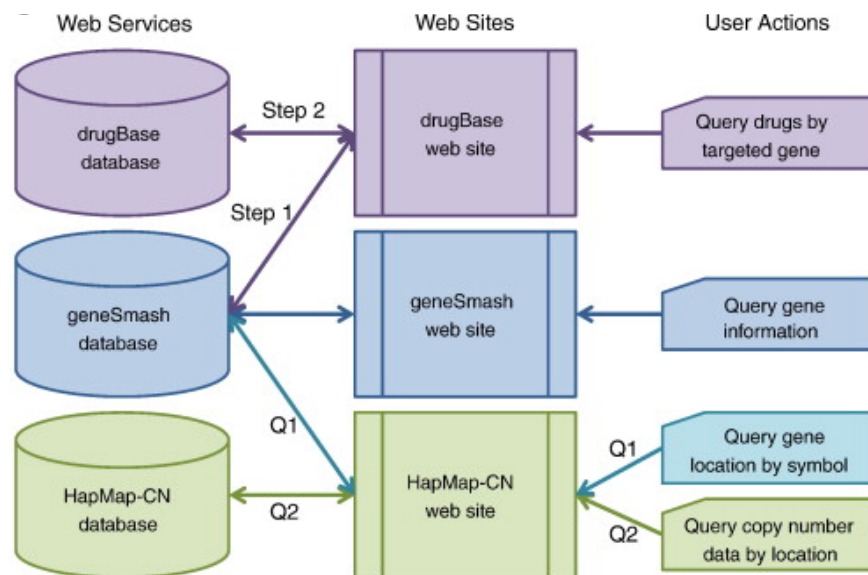
Ensayo - Relax with CouchDB – Into the non-relational DBMS era of bioinformatics

La universidad de Texas junto con el centro MD Anderson Cancer Center, están incursionando con el uso de nuevas tecnologías para facilitar las investigaciones y tener un repositorio de genes facil de mantener y explotar.

Gracias a las nuevas tecnologías como las bases de datos NoSQL, CouchDB que es una base de datos orientadas a documentos, y adoptar arquitecturas web, han creado repositorios de información que son accesibles desde internet, Cuenta con tres aplicaciones que facilitan el acceso a la información:

- geneSmash: http://app1.bioinformatics.mdanderson.org/genesmash/_design/basic/index.html
- drugBase: http://app1.bioinformatics.mdanderson.org/drugbase/_design/basic/index
- HapMap-CN database: <http://bioinformatics.mdanderson.org/HapMap/docs/index.html>

Arquitectura de las aplicaciones



Como se puede apreciar la aplicacion central es geneSmash, que es el repositorio de Genes y tiene relación con las dos aplicaciones drugBase y HapMap-CN.

La adopción de nuevas tecnologías se dio por la complicación de mantener una base de datos relacional, a pesar de que estas bases de datos son probadas y optimizadas por expertos se tiene la desventaja del rendimiento al unir tablas con gran cantidad de filas. Los expertos en bioinformática adoptaron una nueva base de datos orientada a documentos, en este caso

CouchDB al ser orientada a documentos facilita y se adapta perfecto para almacenar datos que no presentan un esquema uniforme como son los genes. Además al ser una base de datos que puede ser accesada con el protocolo HTTP, se elimina el uso de un servidor web y para obtener la información se pueden usar APIs para el manejo de JSON desde HTML. La creación de aplicaciones se vuelve más fáciles y por la eliminación del servidor web son más rápidas de desarrollar.

Conclusión

Hoy en día se cuenta con gran variedad de opciones tanto de almacenamiento de la información como lenguajes para ocupar la misma. Un ejemplo es la adaptación de las bases de datos NoSQL y orientadas a documentos para el área de la genética.

La base de datos NoSQL no serán la única solución ni la definitiva para todos los problemas que nos enfrentemos, los motores relacionales cuentan con mucha historia son los más utilizados y para diferentes problemas son la mejor opción, hay que conocer las ventajas y desventajas de cada una de las opciones para poder tomar la mejor decisión en cada problema a que nos enfrentemos.