

Para poder recomendar películas se ha implementado un algoritmo Map-Reduce que saca usuarios similares. Una vez que se aplica este algoritmo se sacan películas aleatorias que hayan visto algunos de estos usuarios similares. Los usuarios son llamados dentro de nuestra aplicación clientes.

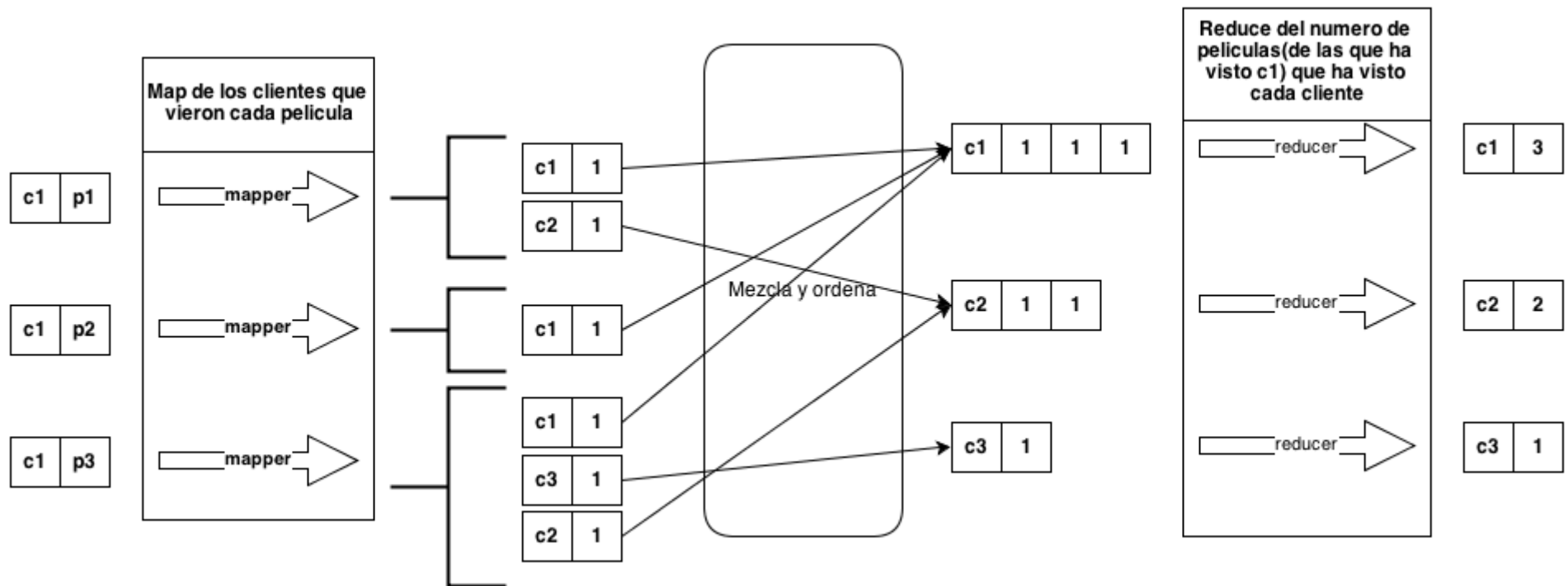
El primer paso del Map-Reduce es sacar todas las películas que han sido vistas por el cliente del que queremos sacar los usuarios similares (que llamaremos a partir de ahora Cliente1). Estas películas las sacamos directamente con una consulta a la base de datos.

Una vez hecho esto tendremos que pasar a mapear cada película para sacar los clientes que ya han visto esa película. Esta parte es la parte Map, que podríamos en un caso hipotético repartir entre varios servidores esclavos, ya que son solamente lecturas sobre los mismos datos (no tiene efectos colaterales). Para cada cliente sacamos una clave del tipo (cliente, 1), siendo 1 la cuenta de que ha visto 1 vez la película.

Ahora tenemos que hacer un mezcla y ordena, para lo cual juntaremos los datos anteriores por clientes, quedando así para cada cliente una lista de 1s que simbolizan cuantas veces se ha visto una película vista por el Cliente1.

Por último debemos de reducir, es decir, pasar toda esa lista de 1s a un valor concreto: simplemente sumar los 1s.

Para ilustrar el map reduce hemos hecho un esquema con un ejemplo para los clientes c1, c2 y c3, y las películas p1, p2 y p3:



Una vez acabo el map reduce, y con esos resultados, tenemos que c1 es el cliente que más concuerda con c1. Esto es lógico ya que es él mismo. Después es c2, que ha visto 2 películas de las 3 que ha visto c1, y por último c3, que solamente ha visto una película de las que ha visto c1. Ahora solamente quedaría sacar una lista de sugerencias de películas, para lo cual sacamos películas al azar los clientes, con mucha más posibilidad en aquellos que más películas en común con c1 tengan.