

One of the largest Price comparison and cashback shopping portal on web

The client

Pepperglobal Ltd is a UK based startup that runs Peppervillage.com, the Internet's comparison shopping and rewards social cooperative. The company is focused on ensuring that expenses around advertising, marketing and staff are kept to the minimum, so that earnings can be redistributed within the member community, assuring every member, "There's Money In It."

Key Business Requirements

One of the key business requirements for Peppervillage.com is to publish millions of shopping content from various diverse advertisers across the globe, categorize, tag those and making them searchable along with multi-store and multi-brand facets in a seamless manner. The portal needs to have a strong affiliate model, where members earn commission for purchases made by friends they bring to the community, as well as for purchases made by their friends' friends on a classical MLM style.

Users should be able to rate, review and tag any shopping and non-shopping content node. Users should also be able to build various content-based social networks with other users of the portal. Portal should offer different role based activities. For example, administrators should be able to do the following:

1. Enable/Disable any functional module
2. Block/Unblock users, moderate any abusive comments or any other content type
3. Set commission and cashback levels for a combination of product category and store
4. Activate/De-activate any shopping node and/or store

Key Challenges

Sheer number of shopping nodes being aggregated from various advertisers across the globe and building a scalable solution around them presented the biggest challenger. At present, the number of nodes is around 30+ million. Second challenge is to make the node data updated in near real time. Merchant catalogues change all the time, and update operations need to happen constantly to make the data up-to-date. Third challenge is to tag all those 30+ million shopping product nodes and make them keyword searchable and distribute it across various facets like brand, store and price. Even some of the best known search engines in the world including don't provide faceted search capability.

Our Solution

Considering the sheer volume of dataset and functionalities around it, it was difficult to find out any existing open source platform that can fulfil our need out of the box. For most of CMS components like user role based registration, login, node review,

ratings etc, Drupal 5 was the platform of choice. The rich core functionality coupled with the actively maintained community contributed modules helped us build lot of rich features quite quickly on top of the Drupal core. A number of existing Drupal modules were used straight out of the box, while some bespoke ones had to be built to deliver the commission and cashback calculations and payouts.

While Drupal provided an ideal platform for lightweight components, for the most of heavy-lifting tasks, we wrote custom non-drupal solutions. Drupal node aggregation module was not good from performance point of view and hence we wrote a pure custom performance enhancing module to load and constantly update millions of node.

For keyword and faceted search, apache solr search engine was the search engine of our choice. The architecture implements APC cache and uses a MySQL database cluster configuration. We achieved a better performance with a master-master replication setup, with separation of automated updates and user initiated writes, as well as read operations.