Search Applications for the Cloud: Future Paradigms

Naren Chittar, nchittar@ebay.com, Principal Engineer, Search Engine Group

Disclaimer: Some of the content in this talk is based on my personal opinion. It does not reflect the views of eBay or my past employers.
Outline

• A short tour of eBay.
• Architecture of search applications.
• Cloud & server side computing.
  – Drawbacks.
• Client side computing.
  – Examples.
  – Architectural challenges.
• Conclusion.
• Q & A.
Welcome to today's online marketplace...

...the market that brings buyers and sellers together in an honest and open environment...

Welcome to eBay's AuctionWeb.

Welcome to our community. I'm glad you found us. AuctionWeb is dedicated to bringing together buyers and sellers in an honest and open marketplace. Here, thanks to our auction format, merchandise will always fetch its market value. And there are plenty of great deals to be found!

Take a look at the listings. There are always several hundred auctions underway, so you're bound to find something interesting.

If you don't find what you like, take a look at our Personal Shopper. It can help you search all the listings. Or, it can keep an eye on new items as they are posted and let you know when something you want appears. If you want to let everyone know what you want, post something on our wanted page.

If you have something to sell, start your auction instantly.

Join our community. Become a registered user. Registered users receive additional benefits such as daily updates and the right to participate in our user feedback forum and the bulletin board.
every 1.6 min. a car is sold

every 1 min. a GPS device is sold

every 7 sec. a pair of shoes is sold
• It is a marketplace or a search engine.

• 100+ million active buyers and sellers worldwide.

• 250 million queries each day to our search engine.

• 350 million items live in more than 50,000 categories.
$68.6 billion in merchandise sold in 2011
• 100 million downloads of eBay mobile app.
• 100 million listings.
• Most expensive vehicle sold: Mercedes-Benz: SLR McLaren, $240,001.
• $10 B expected volume transacted.
• Convenient for auctions.
  • Numbers from http://allthingsd.com/files/2012/09/100million-infographic0371-640x2941.png
Current Search Architecture

Client

User Query

Query

 Ranked Results

Server

Query Processing

Dictionaries

Search

Ranking

Inverted Index
Cloud / Server Side Computing

• Advantages

• Drawbacks.
  – Hardware costs.
  – Network bandwidth.
  – Energy.
  – Cooling.
  – Security.
Future Paradigms

Client

User Query

Some query processing

Intermediate results

Results post processing

Semi processed query

Server

Dictionaries

Search

Inverted Index
Today’s Mobile Client

• Some tasks could be performed on the client.
  – More possibilities on mobile than desktop.

• Compute heavy tasks.
  – iPhone 5 processor
    • Dual core, 1GHz.

• Not memory heavy tasks.
  – iPhone 5 processor
    • Only 1 GB RAM.
    • Many search applications very data centric.
    • need huge data structures ~100s of GBs.
Examples - Text Search

• Query processing
  – Simple regex matching, simple spell check 😊. (iPhone)

• Results Post Processing
  – Clustering (can be O(N^2)).
  – Second or third pass ranking.

• Personalization.
  – Client has more context of user.
Examples - Image Processing and Search

• Compute heavy.

• Good support on mobile.
  – OpenCV for Android and iOS.

• Feature Extraction on client.
  – SURF, SIFT.
  – Saves server side compute and network bandwidth.

• Example Instagram.
  – Minimal server side compute.
  – Server side would make it bankrupt (even quicker).
Examples - Speech and Audio Search

• Compute intensive (Siri, Shazam).

• Good support on mobile.
  – Pocket sphinx.

• Feature extraction on client.
  – Cepstral Features extraction and quantization.
  – Saves server side compute and network bandwidth.
Challenges of Client side computing.

• Updates to client.
  – Change in client algorithms & data.

• Battery life reduction.
  – Algorithm complexity and code efficiency very important.

• Security.
  – Possibility of altering client side algorithms and data.
Conclusion

• Huge potential for client side compute sharing.
• Mobile way ahead.
• Desktops lag behind.
  – Fundamental OS level security enhancements needed.
• Network bandwidth increase will help.
Q & A

Contact

nchittar@ebay.com
nchittar@alumni.stanford.edu