Three Innovations that made Google

HPDC PC lectures at VU Amsterdam
- Friday 03/17/2006

Thomas M. Stricker
Google European Engineering
March 2006
If there is electricity...
… you will find **Google**!

Google Search is available in 109 languages and 114 national and international Internet domains including google.com.
Google: a really **global** brand name
The mission of Google:

Organize all the worlds information and make it universally accessible and usable.
Summary of presentation of today

• The tree most important innovation that lead to Google:
  – **Page Rank** - qui ordonne les résultats selon leur importance - le PageRank.
  – **Targeted advertisements** - fully automated system to place advertisements relevant to the search query based on the mechanism of an **online auction** to set the pricing in any market.
  – **Infrastructure** based on **commodity PCs**, standard networking and entirely fault tolerant software.

• HPDC infrastructures: **map reduce** and **Google File System (GFS)**.
• Introduction to **Google European Engineering Centre** in Zürich
Innovation: PageRank™

Development of Ranking in Web IR

• First search engines:
  – Classic information retrieval techniques
  – Text-only

• Current search engines:
  – Analysis of the web structure
  – Text and link

• Future search engines:
  – (Better) understanding of queries and documents
  – Text, link, entities, concepts, …
Classic Information Retrieval

- **Ranking** is a function of
  - Query term frequency within the document \((tf)\) and
  - Query term frequency across all documents \((df)\)

- **Details:**
  - Assign a weight to each term in a document
  - Document score for query is the sum of the weights of the query terms
  - Rank the documents by decreasing score

- **One approach:**
  
  Weight of term in document = \(f(tf) \cdot idf\),
  
  where \(idf = \log N/df\) and \(N\) is the number of documents in the collection
Web Information Retrieval

• Problems with Web documents:
  – Deliberate misinformation (*spam*)
  – Variety in pages (genre, quality, languages …)
  – Topic mixture

• Goal:
  – Precision in top 10

→ Ranking based only on *query term frequency within the document and across all documents* does not work well, manual identification of high quality documents became important
Web Solution: PageRank™

• Importance:
  – Mostly based on hyperlink analysis

• Assumption:
  A *link* from page A to page B is a *recommendation* of page B by the author of A
  ⇒ Quality of a page is related to number of recommendation
    i.e. its in-degree

• Refinement by recursion:
  Quality of a page is related to
  – Its *in-degree*, and to the *quality* of pages linking to it
  → PageRank [Brin & Page ‘98]
A graphical definition of PageRank™

• Query-independent measure of value for each page
Query [bush]:

- Pages for [bush]: 85’000 000…

Welcome to the White House - [ Traduire cette page ]
Whitehouse.gov is the official web site for the White House and President George W. Bush, the 43rd President of the United States. ...
www.whitehouse.gov/ - 38k - 9 mar 2005 - En cache - Pages similaires

www.whitehouse.gov/president/gwbbio.html
Pages similaires
[ Autres résultats, domaine www.whitehouse.gov ]
Example: query [miserable failure]

• Pages for [miserable failure]: 567'000…

Biography of President George W. Bush - [Traduire cette page]
George W. Bush is the 43rd President of the United States. He was sworn into office on January 20, 2001, re-elected on November 2, 2004. ...
www.whitehouse.gov/president/gwbbio.html - 43k - 9 mar 2005 - En cache - Pages similaires

Welcome to MichaelMoore.com! - [Traduire cette page]
Join our mailing list. ...
www.michaelmoore.com/ - 29k - 9 mar 2005 - En cache - Pages similaires

BBC NEWS | Americas | 'Miserable failure' links to Bush - [Traduire cette page]
'Miserable failure' links to Bush. ... Prank website. Newsday newspaper says as few as 32 web pages with the words "miserable failure" link to the Bush biography. ...
news.bbc.co.uk/2/hi/americas/3298443.stm - 32k - 9 mar 2005 - En cache - Pages similaires
Indexing the Anchor Text

• Anchor text is indexed in addition to the page it points to
  
  <title> My example of a Google “Bomb” <\title>
  <a href="http://www…gov/some_biography.html">Miserable Failure</a>

• Most common anchor text “Click Here” doesn’t help users as much
  
  <a href="http://searchenginewatch.com">Click Here</a> to go to Search Engine Watch for…

• Use a most descriptive anchor text to help the pages you recommend
  
Innovation:
Targeted advertisements and auction based pay-per-click pricing.

Part of a patent portfolio held by Google and Yahoo and mutual agreements.
We focus on **relevance** for our ad products also.
Google AdWords

Ad Ranking = Max Cost Per Click (CPC) x Click Thru Rate (CTR)

- **Maximum CPC:** can be set for an individual keyword or for a group of keywords - pay only when a user clicks
- **Auction model:** set your max CPC bid in AdWords based on how much you value the position
- **Ad discounter:** never pay more than your max CPC bid and usually pay less
- **Optimization:** to improve your ad ranking, you can either increase your CPC bid or work to increase your CTR
Creating Ad Groups and Keywords

- Set your maximum cost-per-click (CPC)
  - Specify the maximum CPC you are willing to pay for the keywords in your Ad Group
  - The Traffic Estimator provides average traffic, cost, and position estimates for each keyword

Google recommended maximum cost-per-click
High response: AdWords rewards **relevant ads**

- Position, or rank, is determined by a combination of **CPC** and **CTR** to ensure relevancy
- Poorly performing keywords are automatically disabled
<table>
<thead>
<tr>
<th>CPC</th>
<th>CTR</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1.25</td>
<td>2.0%</td>
<td>2.50</td>
</tr>
<tr>
<td>$1.50</td>
<td>1.6%</td>
<td>2.40</td>
</tr>
<tr>
<td>$1.50</td>
<td>1.3%</td>
<td>1.95</td>
</tr>
<tr>
<td>$1.25</td>
<td>1.3%</td>
<td>1.62</td>
</tr>
<tr>
<td>$1.25</td>
<td>1.0%</td>
<td>1.25</td>
</tr>
<tr>
<td>$1.20</td>
<td>1.0%</td>
<td>1.20</td>
</tr>
<tr>
<td>$1.10</td>
<td>1.0%</td>
<td>1.10</td>
</tr>
<tr>
<td>$1.00</td>
<td>1.0%</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Better relevance wins better position / makes ad cheaper

Same CPC, increase the CTR by 2x

Increase conversion 25% -- then increase CPC 25%
Innovation:
Infrastructure based on Clusters of Commodity PCs and fault tolerance.


Peak of google.stanford.edu
Why Use Commodity PCs?

• Single high-end 8-way Intel server:
  – IBM eserver xSeries 440
  – 8 2-GHz Xeon, 64 GB RAM, 8 TB of disk
  – $758,000

• Commodity machines:
  – Rack of 88 machines
  – 176 2-GHz Xeons, 176 GB RAM, ~7 TB of disk
  – $278,000

• 1/3X price, 22X CPU, 3X RAM, 1X disk

Sources: racksaver.com, TPC-C performance results, both from late 2002
Hardware architecture

- Internet
- Switch / Router (many Gbps)
- Load Balancer
- 80-node cluster
- 80-node cluster
- 80-node cluster
- 2*1Gbps GigE
Google Servers 1999
Good **infrastructure** is powerful
Lessons Learned in Infrastructure

- Hardware: connectors are a bad idea
- Servers: Simpler designs win
- Servers: Keep everything identical
- Traffic grows faster than you think
- Design everything with “factor N” scalability

Cheap hardware allows more computation per query (search) and more space per user (e-mail).
Fault tolerance in Software

Hardware breaks

• At Google scale, many machines will fail every day
• If you have one computer, it might last three years
• If you have a thousand, expect to lose one a day

Have to deal with failures in software

• Replication and redundancy
  • Needed for capacity anyway

Fault-tolerant software makes cheap hardware practical
Software architecture

Google Web Server

spell checker

ad server

Elapsed time: 0.25s, machines involved: 1000+
Basic **Computing Cluster**

Machine 1
- **Job 0** task
- **Bigmemory job** task
  - GFS Chunkserver
  - Borglet slave
  - GFS Master

Machine N
- **Job 2** task
- **Bigmemory job** task
  - GFS Chunkserver
  - Borglet slave
  - Borg master
Google File System (GFS)

- Master manages metadata
- Data transfers happen directly between clients/chunkservers
- Files broken into chunks (typically 64 MB)
- Chunks triplicated across three machines for safety
GFS Usage at Google

- 30+ Clusters
- Clusters as large as 2000+ chunkservers
- Petabyte-sized filesystems
- 2000+ MB/s sustained read/write load
- All in the presence of HW failures

- More information can be found in SOSP, '03
MapReduce: System Structure

- **Mapreduce binary**
  - insert job
  - wait until done

- **Worker0**
  - Map phase: Generate output keys/values
  - Intermediate files (on local disks)
  - Workers assigned to reduce phase work
  - Sort+Merge duplicate records

- **Input files**
  - split 0
  - split 1
  - split 2
  - Input files split into chunks

- **Worker1**
  - Worker2
  - Worker4
  - Worker5

- **Sstable shard 0**
  - Sstable shard 1

- **Mapreduce master**

---

Google™
MapReduce: Fault-Tolerance

• Programming model makes failure recovery easy
  – If worker W dies, re-execute tasks done by W elsewhere

• Multiple executions to reduce average completion time
  – Spawn backup executions of last few map or reduce tasks
  – Greatly diminishes impact of a few slow machines

• Very robust:
  – Running MapReductions able to tolerate even massive failures
  – e.g. large-scale network maintenance: once lost 1800 out of 2000 MapReduce workers, but MapReduction finished just fine
MapReduce: Uses at Google

- Broad applicability has been pleasant surprise
  - Quality experiments, log analysis, machine translation, ad hoc data processing, ...
  - Production indexing system: rewritten w/ MapReduce
    - ~10 MapReductions, much simpler than old code

- Two week period in 8/2004:
  - ~8,000 MapReduce jobs, >450 different MR operations
  - Read ~1500 TB of input to produce ~150 TB of output
  - ~36,000 machine days, >26,000 worker deaths

“MapReduce: Simplified Data Processing on Large Clusters” to appear in OSDI’04
Google European Engineering Centre

Rationale behind opening an operation in Zürich
Google - is the prototype of a Silicon Valley company.

www.7427466391.com
-> a few mathematical puzzles
-> www.google.com/jobs
The Google European Engineering Centre (v2.0)
Working for Google today

• Engineers: Backbone of the company
• Engineers: Do research, invent and implement
• Productivity: Before everything - other interests are always traded off. (even information security - pay attention but avoid paranoia)
• Management: Flat hierarchy, stays out of the way as much as possible
• Information Culture: Open access to corporate information once you are working for Google.
• Projects/Leads: Googlers hunt in small groups. The catch is often a product.
• Success: Depends strongly on what you do for the company. People who understand the mission and culture will succeed quickly.
You can do a lot with **20%** of your time

**Gregor Johann Mendel**  
**Priest and Abbot**  
*(Discovered the laws of genetics in his 20% time.)*
Important reason for an European Engineering Centre

- You will find there engineers from:
  - Switzerland
  - Germany, Austria and Lichtenstein
  - Sweden, Holland, Italy, Romania, Portugal
  - One or another adventurous American

- Quality of life in Switzerland

  Most of our staff are European Computer Scientists (many with US degrees) that prefer to live in Europe.

- We do recruiting in most European Schools
  - ETHZ and EPFL, Uni Berne, Fribourg, Neuchâtel, Genève, Bâle
  - KTH Stockholm, Chalmers, Imperial College
  - UCLeuven, VU, Uni Saarbrücken
Another **reason** for setting up operations Zürich

---

**e-WorkPermits**

**Work permits for the Canton of Zurich**

Fill out the work permit application step by step online and manage your own data. High levels of data privacy and protection are guaranteed.

Using the electronic form reduces waiting times and gives you a better overview of your application’s status.

Three steps to a work permit:

1. Enter your user name and password on the right.
2. Fill out the electronic form step by step. You can stop at any time and resume filling out the form later.
3. Add the supplementary documents electronically or send them in later by post.

A useful hint: Gather the documents you are likely to need and make sure of the most important facts before you begin. This will help you fill out the form quickly in one session.

**Preparation**

On this site you may only apply for permits for the Canton of Zürich and the cities of Zürich and Winterthur. If you want to work in another canton, please consult the appropriate authority in the canton concerned.

**e-WorkPermits receives Swiss Web Award**

e-WorkPermits, the Canton of Zurich’s application for electronically processed work permits, won the Swiss Web Award in the Public Affairs category.
Google Internet services linked to search

- Alerts
- Answers
- Catalogs
- Directory
- Froogle
- Groups
- Images
- Google Labs
- Local
- Mobile
- News Search
- Scholar
- Special Searches
- University Search
- Web Search
Thank you very much for your attention…

… I will be glad to answer your questions.

tomstr@google.com
Google European Engineering Centre
September 2005