



Processing a Trillion Cells per Mouse Click

Common Sense 13/01
21.3.2013

Alex Hall, Google Zurich

Olaf Bachmann, Robert Buessow, Silviu Ganceanu, Marc Nunkesser

Outline of the Talk

- **AdSpam** team at Google

Why we care about interactive data analysis

- PowerDrill UI: internal **web-app to slice & dice data**
- High-level properties of **PD Serving**
scaling from millions to billions of records
- Key ingredients, comparison to other backends,
algorithmic engineering “tricks”

Google's Keyword Targeted Ads

Search for „trendmode“

AdWords ads / Sponsored links

The image shows a Google search interface for the keyword 'trendmode'. The search bar contains 'trendmode' and the search button is labeled 'Suche'. Below the search bar, there are radio buttons for 'Suche: Das Web', 'Seiten auf Deutsch', and 'Seiten aus der Schweiz'. The search results are displayed under the heading 'Web' and 'Ergebnisse 1 - 10 von ungefähr 217'000 für trendmod'. The results are divided into organic search results and sponsored ads. The organic results include links to 'Mode Trend im Jelmoli', 'Mode Trends-Shops', 'Trendmode', 'Trendmode 24h (Italy Mode)', 'Mode von Vive Maria, Pussy Deluxe, Trendmode, Designermode', 'Trend-Mode & Secondhand-Shop', and 'Trend-Mode, Helga Eggel, Zermatt'. The sponsored ads are highlighted with a purple border and include 'Neu: Mode', 'Trendmode', 'Mode', 'Trendmode de luxe', and 'Trendmode'.

Web Ergebnisse 1 - 10 von ungefähr 217'000 für trendmod

Mode Trend im Jelmoli Anzeigen
www.jelmoli-shop.ch Frische Farben & trendige Akzente - **Mode** für jeden Anlass bestellen.

Mode Trends-Shops Anzeigen
www.bonprixsecure.com/ch/ Der neueste **Trend** zum kleinen Preis jetzt zugreifen und sparen!

Trendmode Anzeigen
www.comelia.ch Aktuelle Damenmode: Chic in Strick Jeans-Shop, Maxi-**Mode** und mehr!

Trendmode 24h (Italy Mode)
Trendmode, Jeans,Capri, Hüftjeans,Crazy Age Jeans,Uncle Sam,Blue Queen Jeans,Kronenshirts,KroneTOPS,Damenjeans,Damenmode,Neckholder, Bekleidung,B*9JOJO ...
www.trendmode-24h.de/ - 30k - [Im Cache](#) - [Ähnliche Seiten](#)

Mode von Vive Maria, Pussy Deluxe, Trendmode, Designermode
Junge Bekleidung für Damen und Herren ist im Angebot. Ein Newsletter ist abonnierbar.
www.dress24.com/ - 7k - [Im Cache](#) - [Ähnliche Seiten](#)

Trend-Mode & Secondhand-Shop
Trend-Mode & Secondhand-Shop, Luzern - ... **Trend-Mode & Secondhand-Shop**. Firmeneigentümer. Meinen Eintrag aktualisieren. Bruchstrasse 46 6003 Luzern ...
www.hotfrog.ch/Firmen/Trend-Mode-Secondhand-Shop - 16k - [Im Cache](#) - [Ähnliche Seiten](#)

Trend-Mode, Helga Eggel, Zermatt
Trend-Mode, Helga Eggel, Zermatt - Betrieb eines Modegeschäftes.
www.moneyhouse.ch/u/trend-mode_helga_eggel_CH-600.1.002.223-7.htm - 22k - [Im Cache](#) - [Ähnliche Seiten](#)

Trend-Mode & Secondhand-Shop Krasnic Slavica, Luzern
Trend-Mode & Secondhand-Shop Krasnic Slavica, Luzern - Betrieb eines Modegeschäftes.

Neu: Mode Anzeigen
Aktuelle **Mode**-Kollektionen für 2008
Gleich bestellen & richtig sparen!
www.Ackermann.ch/Damen**mode**

Trendmode Anzeigen
My Redoute: Noch mehr Fashion
Entdecken Sie Ihren neuen Stil!
www.laredoute.ch

Mode Anzeigen
Die passenden Fashion-Trends für die neue Jahreszeit: bei Quelle.ch!
www.QUELLE.ch/Damen**mode**

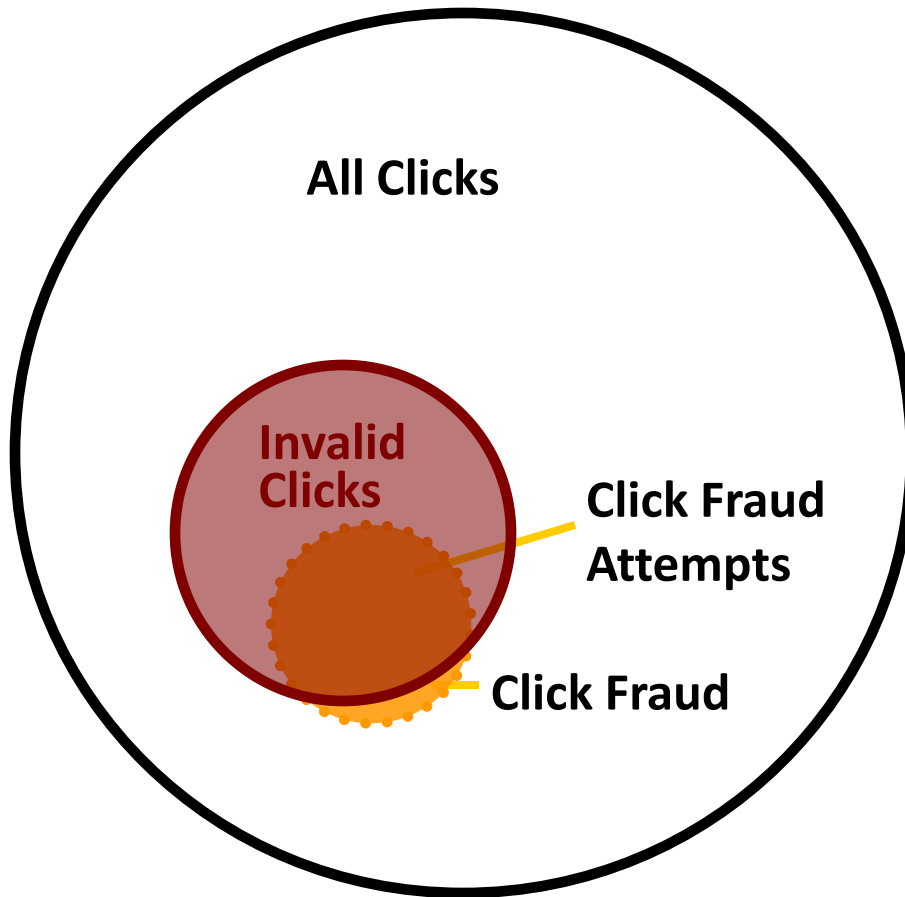
Trendmode de luxe Anzeigen
Feminin. Anspruchsvoll. Exklusiv.
Bestellen Sie stilvolle **Trendmode**.
www.elegance.ch

Trendmode Anzeigen
Mode, Qualität & Stil ab Grösse 42
Preiswert und versandkostenfrei
www.bader.ch

Blusen die Sie anziehen Anzeigen
raffiniert im Schnitt, erstklassig in Qualität u. Verarbeitung
www.daniels-korff.de

- Free search results are independent of ads shown.
- AdSense: earn money by showing Google ads on your own web-page e.g., a cycling page would show adds for bicycle stores

Click Fraud versus Invalid Clicks



Click Fraud

Clicks that are generated with a **malicious or fraudulent** intent.

- Competitor clicking
- Clicking on ads on own page

Invalid Clicks

Clicks on AdWords ads that **Google does not charge** its advertisers for (e.g., double-click on an add).

(Click fraud is only one of several reasons for not charging.)

Note: Diagram is optimized for readability and not drawn to scale.

Google's Multi-Stage Click Fraud Protection

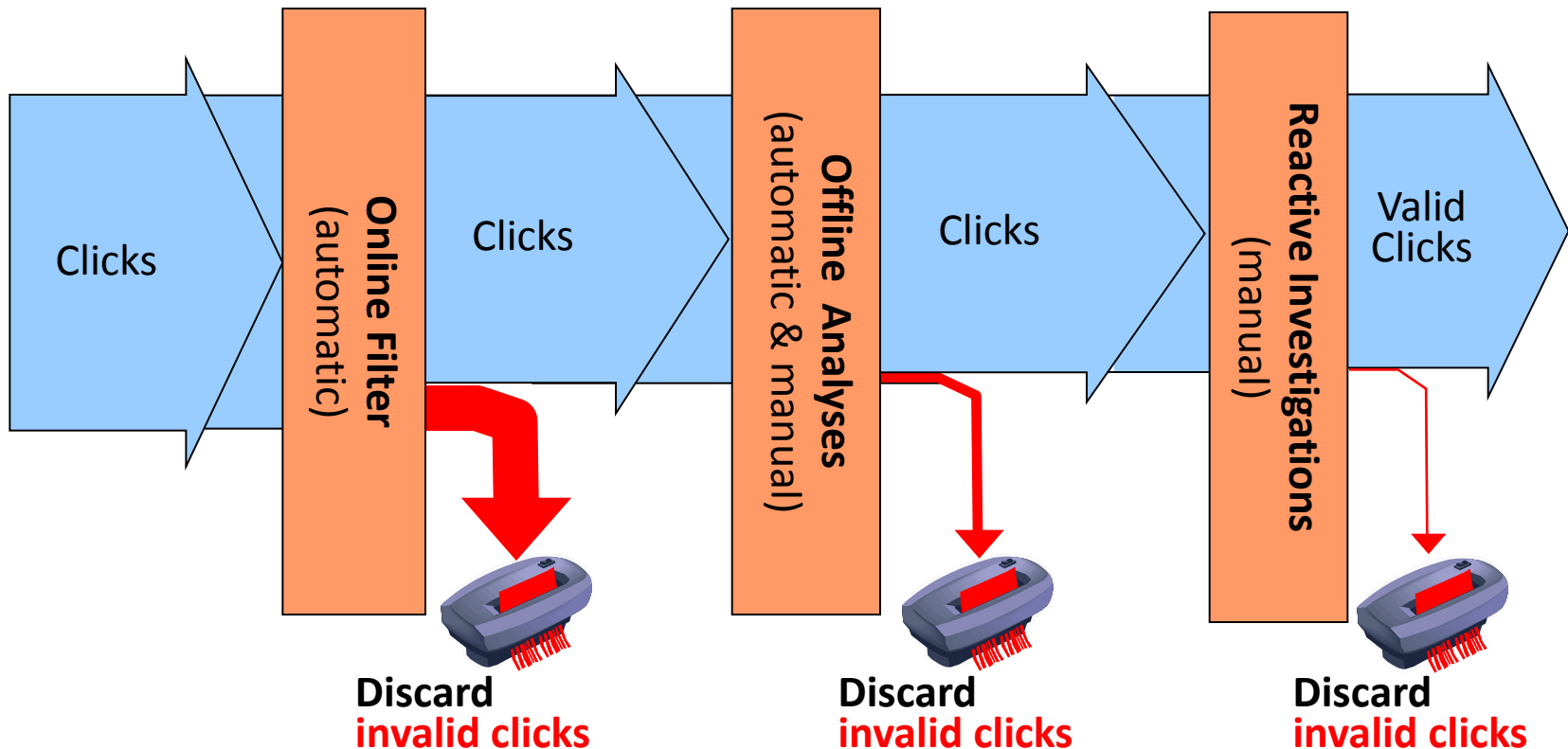
PROACTIVE (<10%)

e.g. double clicks, crawlers, suspicious clicks

e.g. publisher terminations, click ring detection

REACTIVE (<0.02%)

e.g. advertiser credits



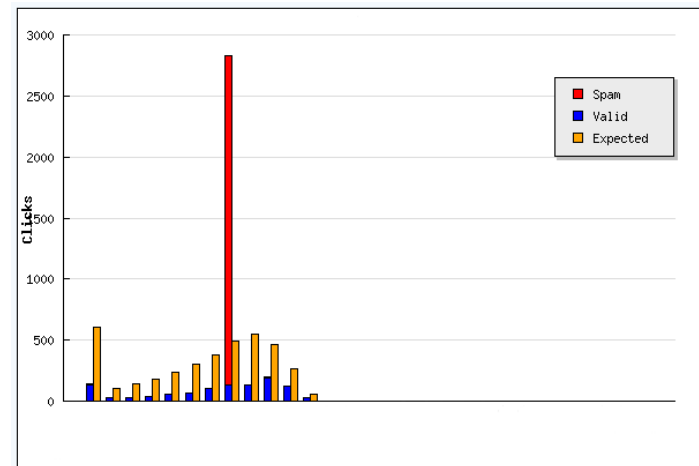
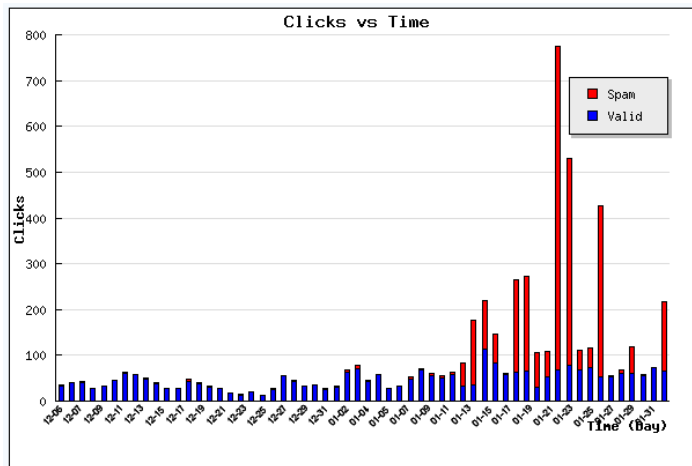
Note: Diagram is not drawn to scale.

AdSpam: Interactive Data Analysis

AdSpam team provides online filters to catch “invalid clicks”

Typical analyses:

- Manually check **set of suspicious clicks**
- **Slice and dice** the data, look at various metrics



Goals:

- **Review:** quickly decide whether clicks are invalid
- **Filter development:** research new filter ideas

PowerDrill UI

Google internal web-app for easy slicing and dicing

- Shows **charts** e.g., clicks over time, top ten countries, ...
- Interactive way of **restricting the data set**

The screenshot displays the PowerDrill web application interface. On the left, there is a sidebar with the PowerDrill logo, a search bar, and sections for 'Defined Fields' and 'Protocol Message'. The main area is divided into several panels. The top panel shows 'Scenario 0' with a 'Where statement' field containing 'Scenario 0 where condition - hint: use the scenario conditions in the charts!'. Below this is the 'Scenario DiffTool' panel, which is currently set to 'By Week'. The central panel features a line chart titled 'Scenario 0 condition' showing 'HoursSpent' (yellow line) and 'Users' (red line) from 2008 to 2012. The chart shows a general upward trend in both metrics. Below the chart is a 'Country' panel displaying a table of the top countries. The table has columns for Country, COUNT, DistinctQueries, Users, and HoursSpent. The data is as follows:

Country	COUNT	DistinctQueries	Users	HoursSpent
	136,728,363	27,787,271	87	53,529.13
US	16,870,870	8,752,600	3,953	21,083.60
IE	11,033,828	5,524,310	293	6,160.00
IN	3,324,987	1,801,822	307	3,029.70
CH	3,104,041	1,448,492	609	5,736.90
Du	140,469	65,718	32	64.60
UK	67,537	31,599	218	599.67
AU	40,309	25,296	147	133.90

PD logs (Google internal queries)

PowerDrill UI

Each chart -> SQL “**GROUP BY**” query
Restriction -> **WHERE** statement

On every interaction

- Send SQL queries to the backend
 CSV, Dremel, PD Serving, RecordIO, ...
- Backend processes SQL on suspicious click data

Needs to be super fast on billions of records!

Column-stores

Basics

- Highly tuned systems to eval SQL queries (analytics)
- Data stored per column, not per row (like most OLTP)
- Access less data per query, fast scans

Commercial Systems

- Hana
 - OLTP and OLAP in one
 - In memory – suggested server-size up to 2TB ram
- Exasol
 - Leader in many of the TPC-H benchmarks
- ...

Column-stores At Google

Dremel

- Streams from disk
- Petabytes of data, millions of tables
- Thousands of light-weight servers
- Fast for 100s of millions of rows

PD Serving **NEW**

- In memory – as much as possible
- Few selected data-sets
- ~1500 servers, 4.5 TB ram
- **Scale to 10s of billions of rows**

wired.com

“Google Crunches One Trillion Pieces of Data With Single Click”

Appeared August '12 www.wired.com/wiredenterprise/2012/08/google-trillion-pieces-of-data/

Contains nice summary of this VLDB article

“Dremel is designed to analyze many different datasets,” says Tomer Shiran, [...], “but this new system is optimized to run in memory, and that means you can achieve really, really low latency.” [...]

“If you have, say, four datasets that are central to your business,” Shiran says, “this is where you would store them.” The system uses various compression techniques, he says, to pack as much data as possible into memory.

Thoughts:

- Obviously, Google cares about more than four datasets
Dremel: disk based, petabytes of data, millions of tables
- OTOH, AdSpam analyses: mostly with two huge (logs) datasets

My Wife...

... after reading the wired.com article wonders about **1 trillion cells per click**:

“This only worked once, right?”

Hmmm, where's the faith?

Reality

- Heavily used within AdSpam since 2 years.
Single user after a “hard day’s work”: up to 12k queries
- Used primarily on 2 major datasets
- Typically a single mouse click triggers 20 SQL queries
- On average these queries process data corresponding to 782 billion cells
i.e., frequently > 1 trillion cells
- Return in 30-40 seconds (under 2 seconds per query)

Remainder of the Talk

- Comparing **existing backends** (latency, mem)
- **Basic data-structures**
- **Optimizations** / algorithmic engineering “tricks”
Stepwise discussion of effects of optimizations
- **Performance** in practice

Comparing Existing Backends

- **CSV files** (comma separated values)
Compute stats by iterating over a csv-file; **scan whole file line-by-line**
- **RecordIO files**
Google binary “record” file-format; **scan whole file record-by-record**
- **Dremel**
 - High performance Google internal column store
 - Columnwise storage: full scan of data, but only necessary columns

	Latency in milliseconds				Memory in KB			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
CSV	55,099	75,207	52,924	71,778	573,339	573,339	573,339	573,339
RecordIO	27,134	50,587	28,497	39,235	551,074	551,074	551,074	551,074
Dremel	7,874	18,191	8,907	48,628	27,943	60,369	118,734	90,792

Basic System

Columnwise storage, per field store:

- **Dictionary**: occurring values \leftrightarrow int “ids”
- Represent the actual data as **list of such ids**

		Latency in milliseconds				Memory in KB			
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Dremel		7,874	18,191	8,907	48,628	27,943	60,369	118,734	90,792
Basic		20	214	179	686	20,001	41,453	132,682	91,232

Q1 Top countries -> 5 mio times counts[countryId]++	Q2 Count & latency / day pre-computed date(..)	Q3 Top tables WHERE restriction many values / ids	Q4 Top tables no WHERE many values / ids
-----------------------------------------------------------	---------------------------------------------------	---------------------------------------------------------	------------------------------------------------

Reduce Memory Footprint

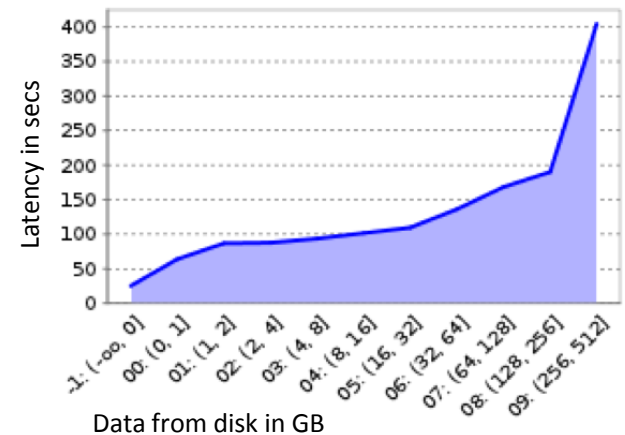
Goal: billions of rows in memory

Compression Savings per Step

	Savings (per step)
Basic (compared to Dremel)	-11% – 34%
Chunks (partitioning)	-16% – 0%
Optimized storage of int ids	11% – 99.5%
Optimized dictionaries (trie)	0% – 78%
Snappy	29% – 49%
Reorder	16% – 55%

Performance

- Latency
 - Reduced from 7-48 **seconds** to 7-260 **milliseconds**
- Memory
 - From **27, 60, 90 MB** down to **35KB, 12MB, 5.6MB**
- **In production, on average**
 - Average response time low # of seconds
 - 92.41% of records skipped
 - 5.02% served from cached results
 - 2.66% scanned
 - 70% of queries fetch no data from disk, 96.5% less than 1GB (overall)



Outlook

Next big topics for our team

- Moving beyond AdSpam
- Cover more use-cases across Google currently at 1700 monthly users
- Go after faster dashboards with < 1s response time