Bring Search Home

Uri Boness (JTeam) & Edwin Adriaansen (ilocal)

JTEAM

ilocal™

About ilocal

- Founded in 2004

- Goal: be the best website for finding companies

- Currently 130 employees

- Surpassed established market players; number one website for local search in the Netherlands

- Expanding our business to other countries

search for... (e.g. Advocaat) restaurant
in or close to... (e.g. 1012XZ) amsterdam
SEARCH
search on category

local :: log in ▾

Search box with spelling suggestion, history etc.

| Results | Map view |

searchterm: restaurant > Restaurants   location: Noord-Holland > Amsterdam

Local results for restaurant in Amsterdam

results 1 to 10 from **2,465** | next

Refine search
Refine on distance
From centre Amsterdam:
within 0.5 km (328)
within 1 km (1,021)
within 2 km (1,644)
within 3 km (2,015)

Verfijn op wijk
Binnenstad (1194)
Oud-Zuid (411)
Oud-West (148)
more ▾

Refine within Restaurants (2286)
Specialism
Dutch restaurants (622)
Italian restaurants (264)
Chinese Indonesian.. (198)
more ▾
Services
Take away meals (363)
Terrace (124)
Private parties or.. (102)
more ▾
Products
Lunch (353)
Dinner (254)
Shawarma (140)
more ▾
Facilities
Internet (171)
Pub (161)
Patio (107)
more ▾

Facet search

Paid vs organic results

Topresultaten

Restaurant Damrak
Damrak 49, 1012 ... Amsterdam
Category: Restaurants
more info | call me now | phone | e-mail | website | route | send to mobile
0.4 km   write review

Restaurant de Korenaer
Damrak 50-51, 1012 LL Amsterdam
Category: Restaurants
more info | call me now | phone | e-mail | website | route | send to mobile
0.4 km   write review

Alle resultaten gesorteerd op: **relevance** | distance | alphabet

1 Restaurant Wissenkerke Sloterdijk
Radarweg 236, 1043 NV Amsterdam
Category: Restaurants
Door Fred (8 april 2008): lekker eten. Heb daar heerlijk gegeten voor...
more info | phone | e-mail | website | route | send to mobile
4.4 km   1 review

2 Yamazato Restaurant
Ferdinand Bolstraat 333, 1072 LH Amsterdam
Category: Restaurants
Door review82 (5 mei 2008): Japans eten op z'n best. Al jaren de ...
more info | phone | e-mail | website | route | send to mobile

3 Restaurant "Manolo"
Warmoesstraat 13/HS, 1012 HT Amsterdam
Category: Restaurants
Door USreviewer (5 maart 2009): nice restaurant, excellent dutch food
0.6 km   1 review

Restaurant Osaka
...straat 22a
1025 KN Amsterdam
2 reviews
Category: Restaurants
more info | website | phone | route

map

show ☐ ATM ☐ parking garages
Location: **center Amsterdam** change

CHEAP AND EASY
☑ BELLEN TEGEN DE VOORDELIGE TELE2 TARIEVEN
☑ RAZENDSNEL INTERNET TOT 20 MB/S

ilocal

JTEAM

How did we start

- We outsourced most of the software development

- A commercial search engine (known by the founder) was licensed

- Learning from users experience

Why Commercial Software

- Off the shelf (quick start)

- Proved to be successful

- Maintenance organization in place

The problems

- We learned faster than our supplier

- Time to market took too long

- Expensive (considering the expansion to other countries)

- Our ideas became standard available in their products

The Alternatives

- Another Commercial Search Engine

- Building it from scratch ourselves

- Use existing open source search engine and modify it according our needs

What did we need

- Scalable solution

- Performance

- GEO locations

- Various ranking strategies

- Control of language specific items

- Knowledge continuation

Project Approach

- Identify technical risks

- Phase 1: Build the search engine

- Phase 2: Enhance the web site to support search engine

- Deployment

- Aftercare

Achievements

- Project delivered in time

- No drop in user visits during migration

- After 2 weeks, the system was stable (we suffered minor bugs and memory leaks)

- After 4 weeks of measurement the average time needed to execute a query dropped from 200ms to 35 ms

Bring Back to Community

- ## The eternal dilemma:

  - Bringing sources back would mean that we make our business plans public

  - It certainly would be appreciated by our competitors

- ## Solution

  - First only contribute generic improvements
  - Later releases: contribute the 'previous release'

Conclusion

- Open Source made ilocal less depended and more competitive

- Lucene / Solr proved to be stable

- Huge performance improvement

- Improvements are easy to implement

- From Business Requirements

- To Technical Implementation

Challenges

- Complex ranking schemes
  - Context aware
  - Dynamic
  - Configurable

- Geo-Location search
  - Sort/Rank by distance + radius facets

- Multi-lingual support
  - Fine grained tuning per language
  - Context aware

- Performance!!!!

Overview

- (Very) Quick overview of Solr Architecture

- Ranking schemes

- Geo Location search

- Multi-lingual support

- Data Import handler

- Solr Multi-core

- Performance & Scalability

- Development process

Solr Architecture

- Schema

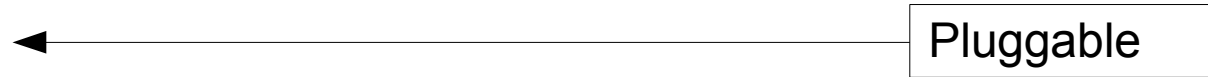 - Field Types ← Pluggable

  - Analyzers & Token Filters ← Pluggable

- Request Handlers ← Pluggable

- Search Components ← Pluggable

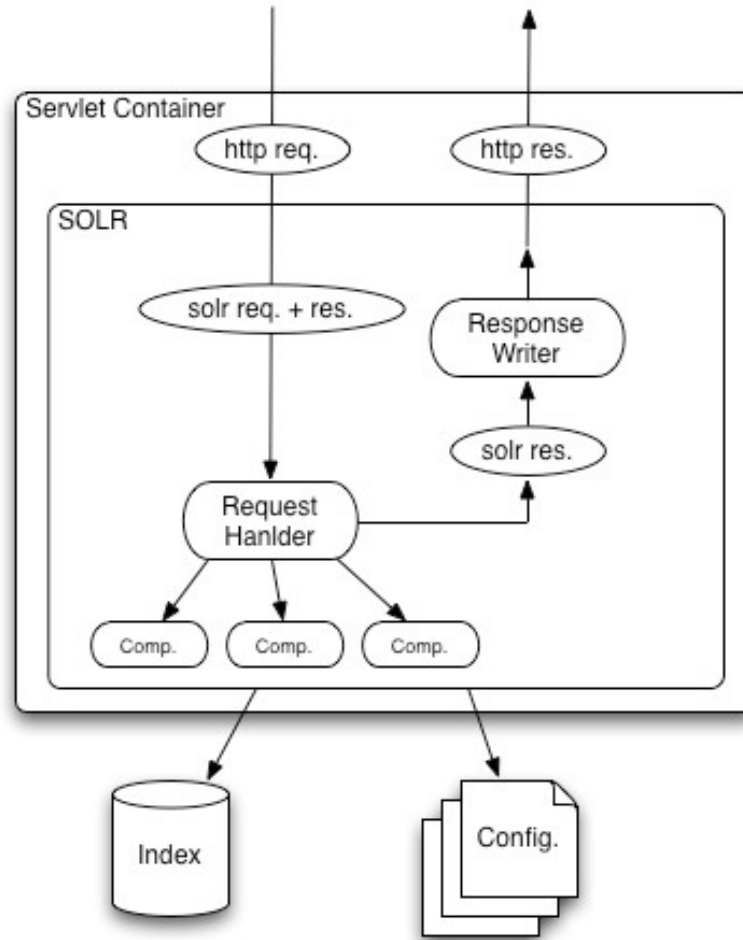- Query Parsers (Factories) ← Pluggable

- Caching ← Pluggable

Solr Architecture Overview

Ranking Schemes

- Introducing the SearchContext

- Ranking Scheme

  - Uses `DisMaxQParserPlugin`

  - Uses Function Queries

- **`IlocalRequestHandler`**

  - Based on the `StandardRequestHandler`

  - Resolves the appropriate search context for each request

Geo-Location Search

- Based on local-lucene & local-solr libraries

- LocalSolrSearchComponent

  - Replaces the default search component

  - Collects geo-location data while searching

    – Longitude / Latitude

  - Filters documents based on geo-location filters

    – Radius list

  - Sorts documents based on distances

  - Utilizes multi-core processor (java 1.5 concurrency)

Multi-lingual Support

- It's all about field types

  - A dedicated field type per language

    - text_en, text_nl, etc..

  - Use language specific analyzers

- The query language is part of the SearchContext

- After resolving the language, the appropriate search field is determined

Data Import Handler

- Provides out of the box data import functionality
  - Databases
  - HTTP (HTML/XML)
  - File System
- Fairly flexible in terms of customization
- We did need to tweak it a bit to use JDBC DataSources
- Bottom line: in just a couple of days we had a full data import functionality ready.

Solr Multi-Core

- Used to manage multiple separate indexes

- In iLocal we used two

  - Companies

  - Locations

- Advantages:

  - Smaller compact indexes perform better

  - Finer control over index configuration

  - Rebuilding each index separately

- Played a major role in our clustering solution

Performance

- Use caching where possible

  - Solr pluggable caching mechanism

- Have fine control on what components are executed

- Don't be scared to write your own customized components

- Bottom line: Dropped average search time in about 70%!!!

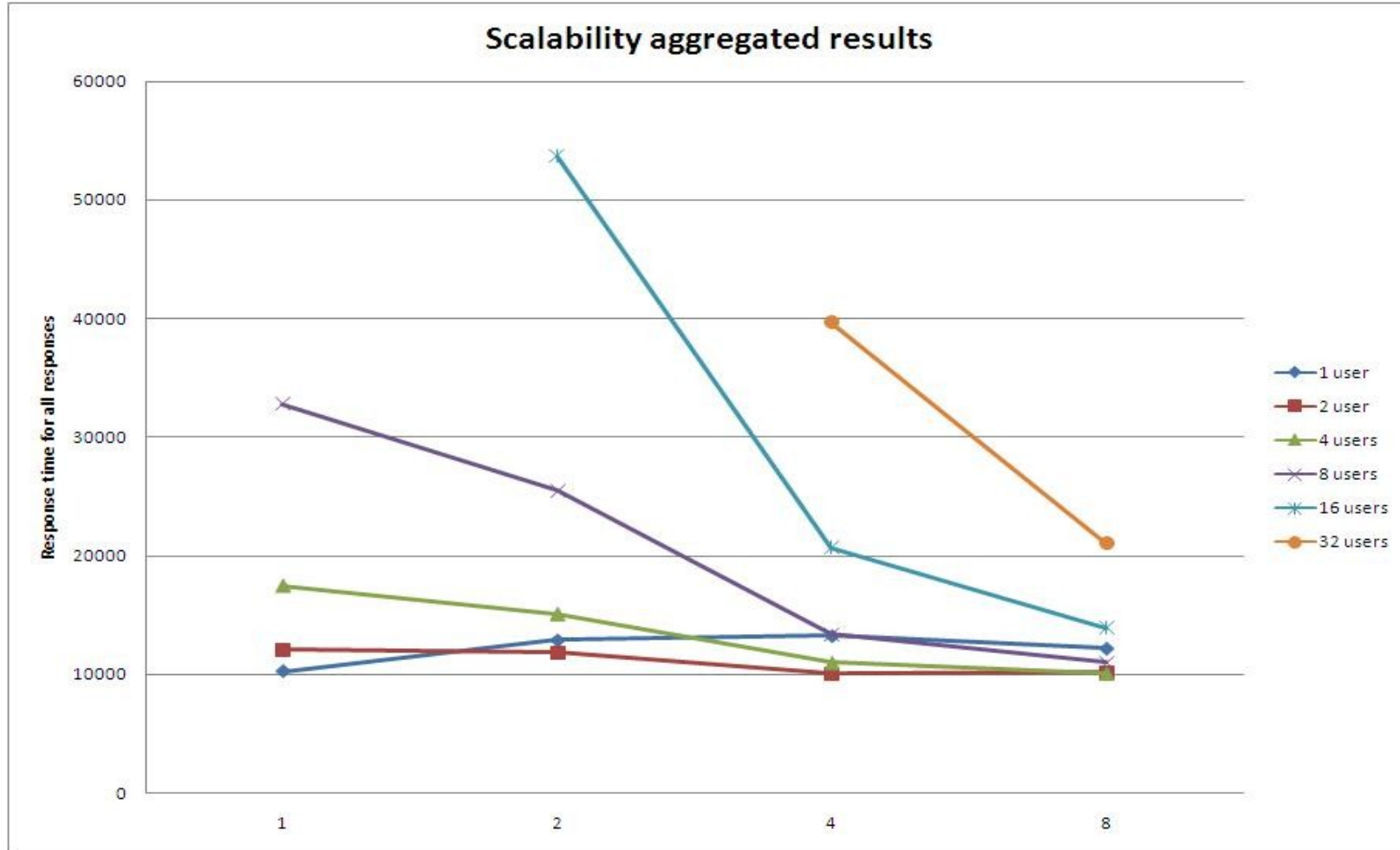- The most complex query doesn't exceed 500ms

Scalability

- The update process
  - Nightly feeds
  - Complete rebuild of the index

- Solr 1.3 enables replication on the index itself
  - No replication for configuration files
  - Fixed in 1.4

- Stateless
  - Ideal for load balancing
  - Scale out by just adding machines

Scalability

Development Process

- Extended Solr Test Harness

- Using JMeter to load test

- Nightly build and deploy

  - The source is checked out

  - Solr is redeployed from scratch

  - Data Import is triggered

  - Tests are ran over the new instance

- Custom GWT based client for user acceptance tests

- Amazon EC2 for scalability tests

Conclusion

- Solr embraces customization

- Feature rich

- Fast! Fast! Fast!

- Ready for the enterprise
  - Scalable

- Low cost!

- Solr & Lucene do the Job!

- **Q & A**