

# GeoCouch: A spatial index for CouchDB

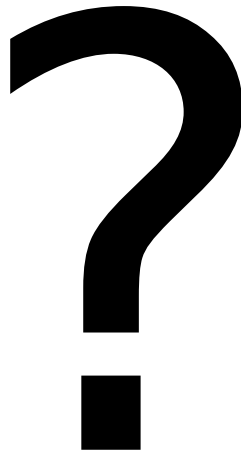
Volker Mische

Free and Open Source Software for Geospatial  
September 6–9th 2010  
Barcelona

# What is GeoCouch?

---

A spatial index for CouchDB



# NoSQL

# Documents

---

```
{
  "_id": "Station-2942_001",
  "_rev": "3-77f17a55f6ab11f7f6668e63a75f2281",

  "name": "Station-2942",
  "date": "2010-09-09",
  "loc": [2.183333, 41.383333],
  "country": "Spain",
  "temperature": 25,
  "rainfall": 3
}
```

# RESTful HTTP API

---

The four basic functions of a database (CRUD):

Create HTTP **PUT** /db/Station-2942\_001

Read HTTP **GET** /db/Station-2942\_001

Update HTTP **PUT** /db/Station-2942\_001

Delete HTTP **DELETE** /db/Station-2942\_001

# Fancy Features

# Schema free



# Arbitrary input/output

---

JSON + JavaScript  
=  
anything

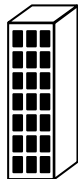
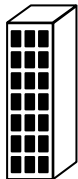
# Concurrency/Robustness

---

- No lock on reads (MVCC)
- Append-only
- Durability (crash-only design)

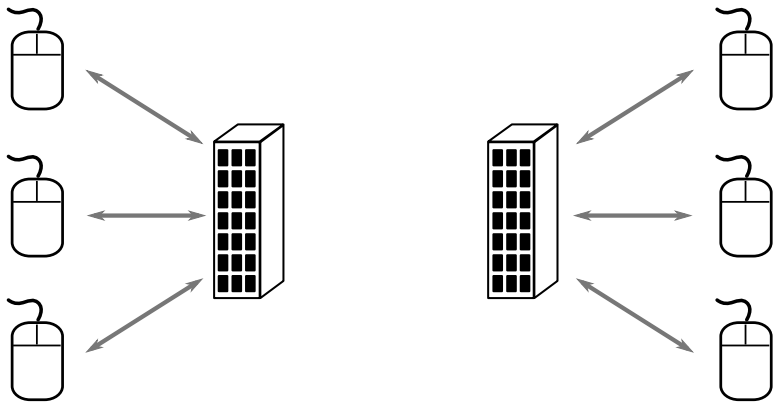
# Replication

---



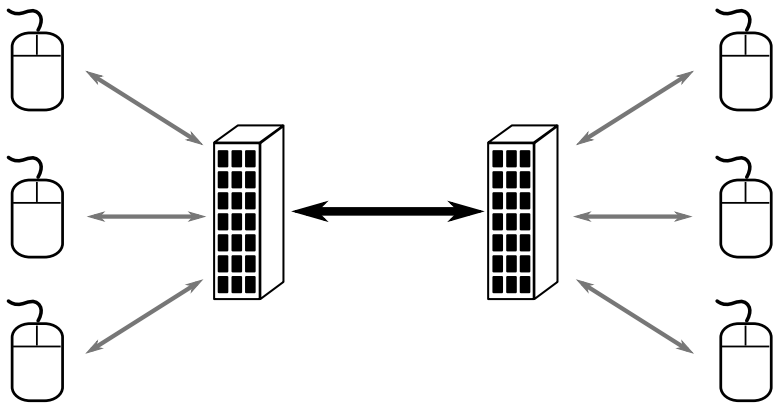
# Replication

---



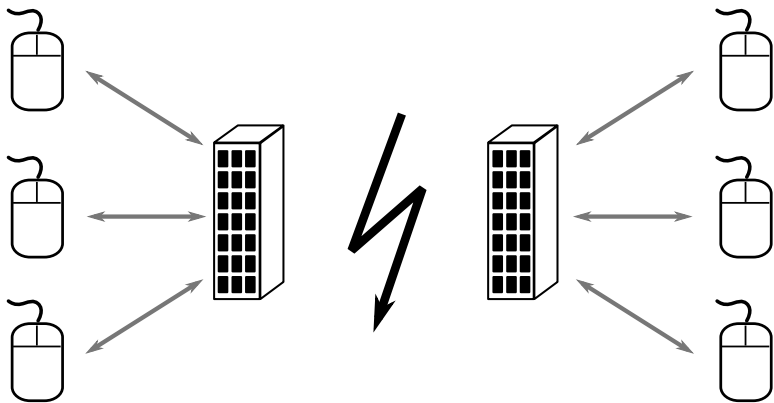
# Replication

---



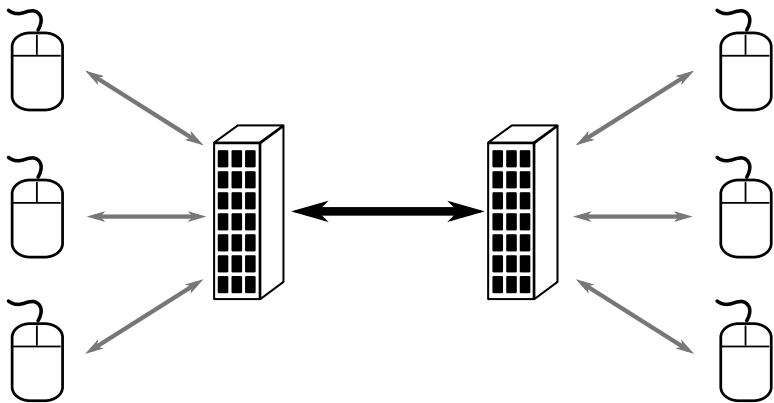
# Replication

---



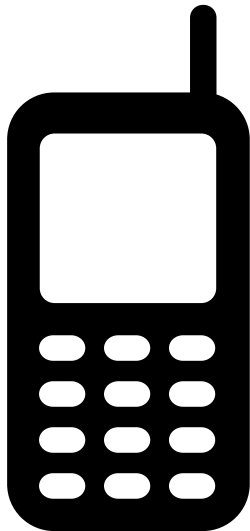
# Replication

---



# Replication

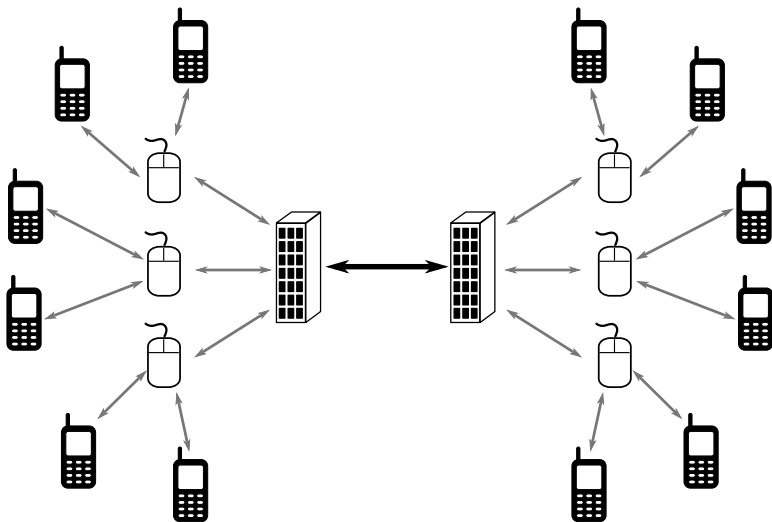
---





# Replication

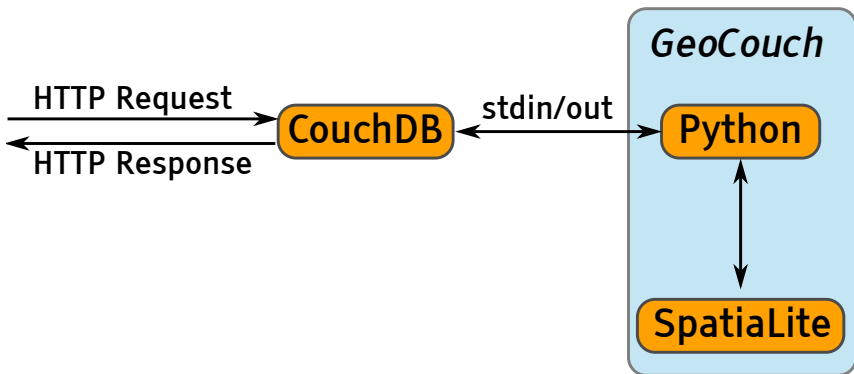
---



# GeoCouch

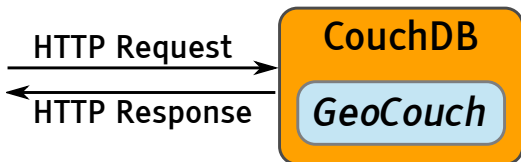
# Previous version

---



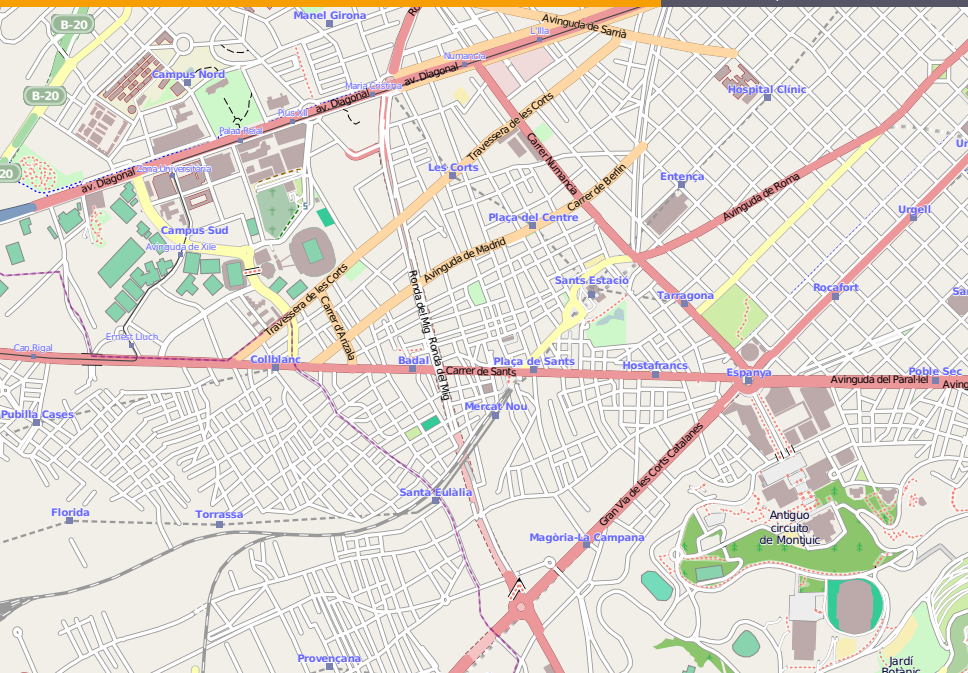
# Current version

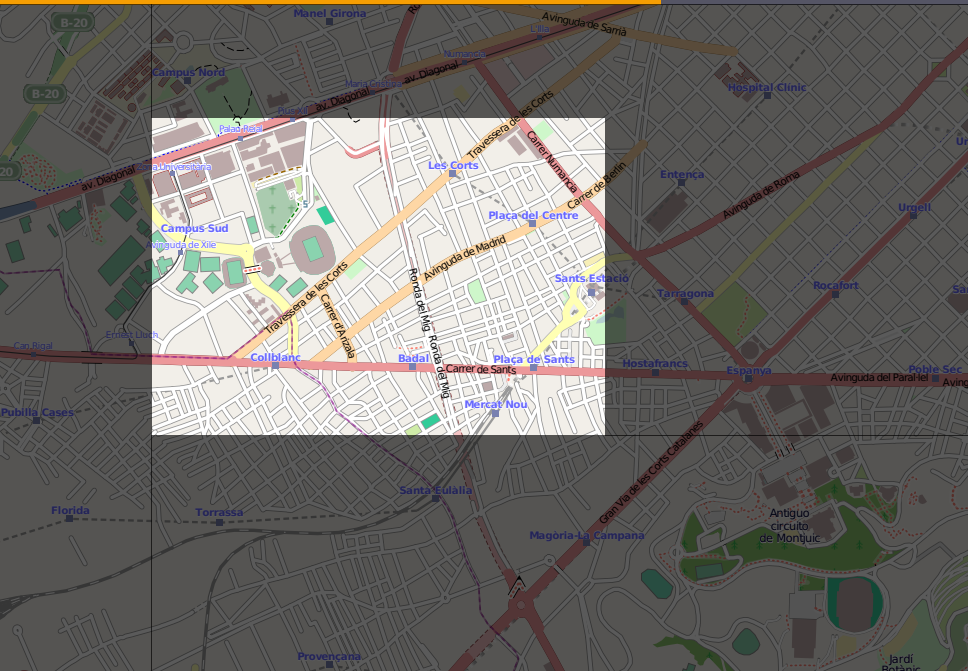
---



# GeoCouch Features

**P**oint  
**P**olygon  
**L**ineString  
**M**ultiPoint  
**M**ultiPolygon  
**M**ultiLineString  
**G**eometryCollection







# Count

# List function (arbitrary output)

# R-Tree vs. Geohash

---

	<b>R-Tree</b>	<b>Geohash</b>
Data Structure	R-Tree	B-Tree
Approximation	bounding box	point
BBox precision	exact	approximate
Good for	all geometries	points
Used by	PostGIS, SpatiaLite, Oracle Spatial	MongoDB

# Benchmarks

# Benchmarks: Competitors

---

- PostGIS 1.4.0
- SpatiaLite 2.4.0RC2
- GeoCouch tag: diplomathesis

# Benchmarks: Datasets

---

100 000 points



USGS GNIS

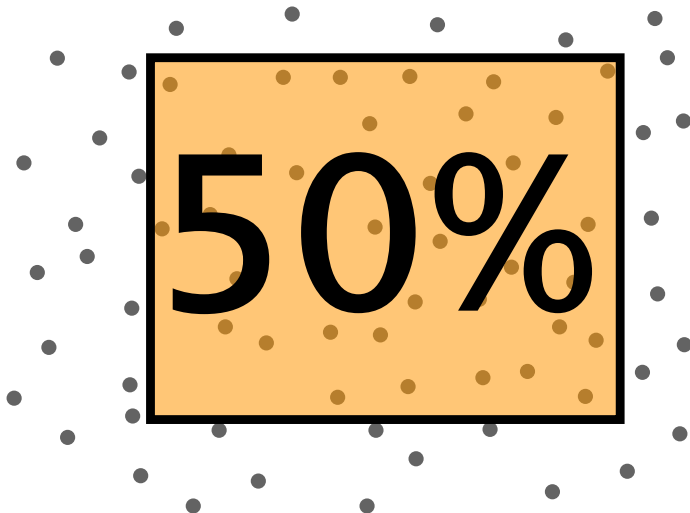
350 000 polygons



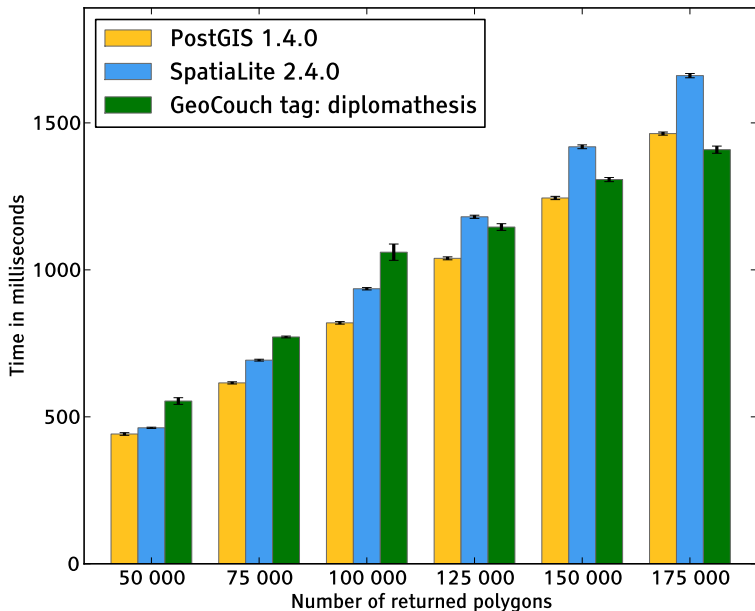
TIGER

# Benchmarks: Test 1

---

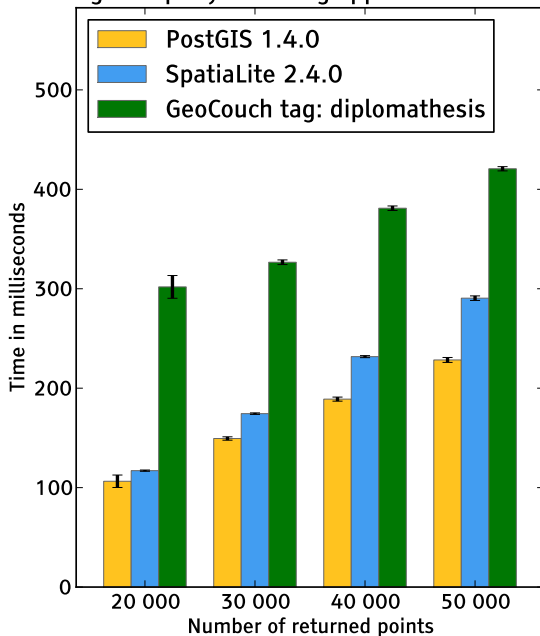


## Bounding box query returning approx. 50% of the polygons



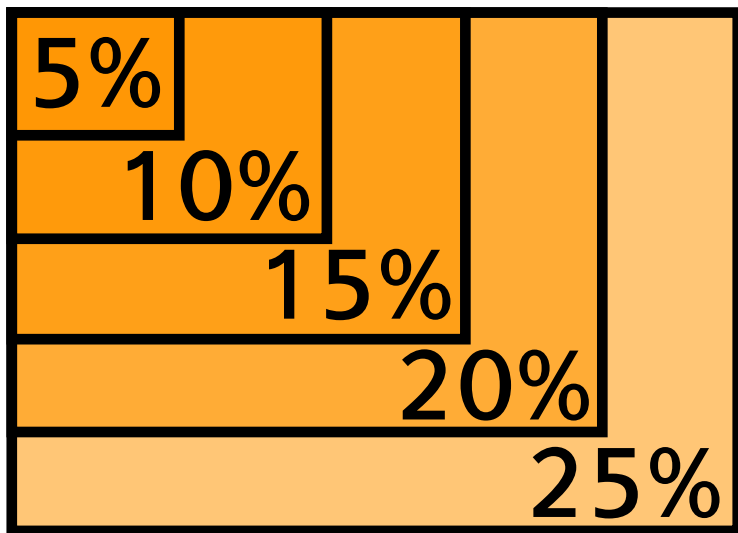


## Bounding box query returning approx. 50% of the points

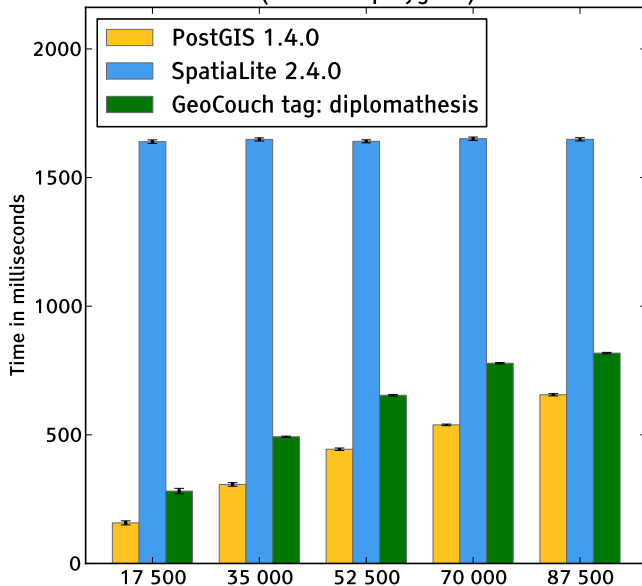


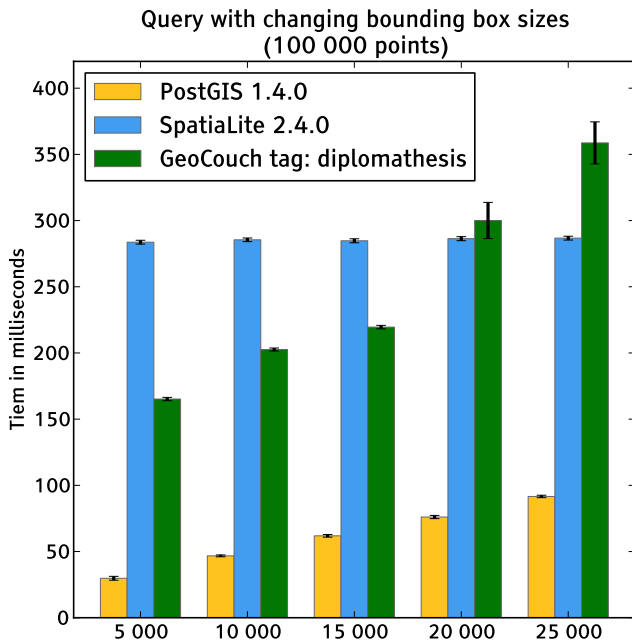
## Benchmarks: Test 2

---



## Query with changing bounding box sizes (350 000 polygons)





# Applications



## PDX API

UPDATE 7/7/10: Coming soon: **Portland Food Cart Finder iPhone App**, powered by PDXAPI! Available for free in the App Store soon. Follow my tweeters if you want to stay up to date: [@maxogden](#)

PDXAPI is a JSON API that provides access to data from [CivicApps](#), the open data initiative by the City of Portland.

Much of the data on CivicApps is in a very raw and hard to use form. PDX API exists to filter the data from CivicApps and provide it in the easiest way possible to application developers wanting to consume the data.

### API

[pdxapi.com/resource?bbox=x0,y0,x1,y1](#)

You must specify the [resource](#) (dataset) and supply a bounding box. The bounding box pair, [x](#) and [y](#), represent the lower left and upper right points of the bounding box, respectively.

You'll need to request the data using library that supports [JSONP](#), such as [jQuery](#).

*Drag the map around to explore bike rack locations*

### Example

**jQuery** JSONP request to the [bicycle\\_parking\\_pdx](#) resource (bicycle lock rack locations):

```
$.ajax({
  url: 'http://www.pdxapi.com/bicycle_parking_pdx',
  dataType: 'jsonp',
  data: {
    "bbox": "-122.6785969,45.5167974,-122.6763439,45.51772964"
  },
  success: function(data){
    // manipulate data
  }
});
```

### Datasets

Datasets currently available through PDX API:

```
bicycle_parking_pdx
bridges_pdx
business_associations
buslines
busstops
cities
cityhall
co_fill
council
counties_pdx
...
```

<http://pdxapi.com/> - spatial index enabled flavor of CouchDB. This means that every piece of data in PDXAPI is accessible and

Awesome

http://localhost:3000/account

gotAPI UF + = delicious v . Transport Projects Services v [ + calagator ] Cornify ☒ pastefire

Aweso... Aweso... http://... http://... has\_m... http://... http://... bbox o... Googl... styles... git justin... HTTP

GeoNotifications for reidab@gmail.com

[Edit User](#) | [Logout](#)

## Contact Methods

- Phone Call to 15038302651  
[\[Edit\]](#) [\[Delete\]](#) [\[Test\]](#)
- "orange", a Text Message to 15038302651  
[\[Edit\]](#) [\[Delete\]](#) [\[Test\]](#)
- "Chris", a Phone Call to 7342761100  
[\[Edit\]](#) [\[Delete\]](#) [\[Test\]](#)
- [Add a contact method](#)

## Schedules

- Bike Parking PDX**  
Subscription:  
Area:
- Contact Method

## Your areas of interest

Map data ©2010 Google

# Next steps

---

- Polygon search (→ Erlang geometry library)
- Sharding
- Layer on top (e.g. for clustering, reprojection)



# Thanks!

# Get in touch with me

---

- Webiste: <http://vmx.cx/>
- IRC: vmx @ freenode
- Email: [volker@couch.io](mailto:volker@couch.io)
- Jabber: [volker@vmx.cx](xmpp:volker@vmx.cx)