CouchDB

Boyoung Yoo, Nick LaRosa, Michael Powers
Goals

- Measure speedup of a distributed web application using Amazon Web Services
CouchDB Model

RDBMS

Customers Table:

<table>
<thead>
<tr>
<th>Name</th>
<th>Id</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate Y</td>
<td>1</td>
</tr>
<tr>
<td>Small Biz Z</td>
<td>2</td>
</tr>
</tbody>
</table>

Orders Table:

<table>
<thead>
<tr>
<th>ID</th>
<th>Customer Id</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Content Management System</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Search Engine Optimization</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Donation System</td>
</tr>
</tbody>
</table>

CouchDB

Type: Customer
Name: Candidate Y
Orders: [Donation System]

Type: Customer
Name: Small Business Z
Orders: [S.E.O, C.M.S.]

Consistency
Availability
Partition tolerance
Eventual consistency
Consensus protocols for HA consistency
Enforced consistency
Paxos
Dataset

Apollo
Toward Fact-finding for human-centric sensing

Overview
People
Publications
Dataset

Dataset Overview

We put here several datasets being used in some of our demos. Please cite the source of these datasets as the 'Apollo Project, Department of Computer Science and Engineering, University of North Carolina' when using these datasets in your publications, demos, or presentations.

Please email kogelnik@cs.unc.edu to contact for collaborations or request username/password to download these datasets. Also, if you have a request for data about an on-going event, please suggest via email as well.

Crimea Crisis

Around 1.9 million tweets were collected from February 17, 2014 to April 9, 2014. [Download]

Syria Tactical Weapon

Around 206,000 tweets were collected from August 22, 2013 to August 31, 2013. [Download]

Boston Marathon Bombings

Around 44,500 tweets were collected from April 15, 2013 to April 16, 2013. [Download]

Hurricane Sandy

Around 404,300 tweets were collected from October 27, 2012 to November 16, 2012. [Download]
Current Configuration

Presentation Tier → Logic Tier → Data Tier

1. Twitter Investigator
2. RESTful Service Port 5984
RESTful API

Operation

GET /search

Request

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>database</td>
<td>string</td>
</tr>
<tr>
<td>query</td>
<td>string</td>
</tr>
</tbody>
</table>
## RESTful API

### Operation

**PUT /mark**

```json
{
  "ID": string
}
```

### Request

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>database</td>
<td>string</td>
</tr>
<tr>
<td>tweetList</td>
<td>array of tweet objects</td>
</tr>
</tbody>
</table>
RESTful API

Operation

PUT /highlight

Request

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>database</td>
<td>string</td>
</tr>
<tr>
<td>tweetID</td>
<td>string</td>
</tr>
<tr>
<td>text</td>
<td>string</td>
</tr>
</tbody>
</table>
Test Cases

● **Control Case:** Single Server + DB Instance

● **Testing Methodology:**
  ○ Condor jobs for 100 and 1000 clients
  ○ Record rates of 1, 10, 50, 100, 1000 requests per second via ELB Access Log File