CS122D: Beyond SQL Data Management
—Lecture #12 —

Mike Carey
UC Irvine
mjcarey@ics.uci.edu

M. Carey, Spring 2021: CS122D
Announcements

• You’re over half-way there...!

| W 4/21  | Document stores: JSON and MongoDB | Ch. 9 NoSQL Distilled (old!) |
| M 4/26  | Document stores: MongoDB (cont.)  | MongoDB materials (as needed) |
| W 4/28  | NoSQL DB design principles        | Ch. 3 NoSQL Distilled         |
| M 5/03  | **Midterm Exam (Checkpoint)**     | **3:30-4:50 PM -- be there!!!** |
| W 5/05  | Document DBs: Couchbase Server & N1QL | CB Analytics paper |
| M 5/10  | Document DBs: Couchbase Server (cont.) | CB materials (as needed) |
| W 5/12  | Graph DBs: Graph modeling & Neo4J | Ch. 11 NoSQL Distilled (old!) |
| M 5/17  | Graph DBs: Neo4J (cont.)          | Neo4J materials (as needed)   |
| W 5/19  | Big Data Analytics: Google, MapReduce, HDFS | Big Data Platforms paper (skim) |

• HW #4 should be in progress...

| HW4     | Th 5/06 | Mo 5/17 (11:59 PM) | CB Server | HW4 Setup | HW4 Details ➔ Template | HW4 Solution |
| HW5     | Mo 5/17 | Th 5/27 (11:59 PM) | Neo4J     | HW5 Setup | HW5 Details             | HW5 Solution |

• **Today:** *A N1QL for your thoughts*...
But First: Consistency in CB Server

• Write options (durabilityLevel)
  • Default — write to memory on active node
  • Majority — write to memory on majority of replicas
  • MajorityAndPersistToActive — Majority, plus write to disk on active node
  • PersistToMajority — Majority, plus write to disk on majority of replicas

• Read options (get)
  • Default — read from active node
  • replica=True — if active node unavailable, read from replica (for availability, trading off consistency)

• Query options (index consistency for GSI use)
  • Not_bounded (default) — read from active node
  • At_plus — indexes updated to logical timestamp of application’s last update
  • Request_plus — indexes updated to query’s timestamp
N1QL – Resource Reminders

• Some useful online N1QL tutorial material:
  • https://query-tutorial.couchbase.com/tutorial/
  • https://sqlplusplus-tutorial.couchbase.com/
  • https://docs.couchbase.com/server/current/analytics/primer-beer.html

• Also see:

Donald D. Chamberlin

American computer scientist

Donald D. Chamberlin is an American computer scientist who is best known as one of the principal designers of the original SQL language specification with Raymond Boyce. He also made significant contributions to the development of XQuery.
Clarification: Query Service Indexing

```sql
SELECT * FROM couchmusic
WHERE type = "userprofile";
```

```
[
{
  "code": 4000,
  "msg": "No index available on keyspace couchmusic that matches your query. Use CREATE INDEX or CREATE PRIMARY INDEX to create an index, or check that your expected index is online.",
  "query": "SELECT * FROM couchmusic WHERE type = "userprofile";"
}
]
```

```sql
CREATE PRIMARY INDEX ON couchmusic;
CREATE INDEX adv_type ON couchmusic(type);
```

Q: Tradeoffs...?
Indexing – N1QL for Query

• Similar to physical design for a relational DB

Query History

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 10 | DROP INDEX orders.adv_ship_date;
|    | ✓ success |
| 11 | SELECT VALUE c FROM customers c WHERE c.address.zipcode LIKE "6%";
|    | ✓ success |
| 12 | CREATE INDEX adv_address_zipcode ON `customers`(`address`.`zipcode`)
|    | ✓ success |
| 13 | SELECT VALUE cust FROM customers cust WHERE cust.address.zipcode LIKE "6%";
|    | ✓ success , 3 documents |
| 14 | DROP INDEX customers.adv_address_zipcode
|    | ✓ success |
Covering Indexes – N1QL for Query

• Goal: Keep bigger queries away from the (KV) data!

<table>
<thead>
<tr>
<th>Query History</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>filter queries</td>
<td></td>
</tr>
<tr>
<td>5 CREATE PRIMARY INDEX ON orders;</td>
<td>✓ success</td>
</tr>
<tr>
<td>6 SELECT VALUE o FROM orders o;</td>
<td>✓ success, 8 documents</td>
</tr>
<tr>
<td>7 SELECT MAX(o.ship_date) AS most_recent FROM orders o;</td>
<td>✓ success, 1 documents</td>
</tr>
<tr>
<td>8 CREATE INDEX adv_ship_date ON <code>orders</code>(<code>ship_date</code>)</td>
<td>✓ success</td>
</tr>
<tr>
<td>9 SELECT VALUE MAX(o.ship_date) FROM orders o;</td>
<td>✓ explain success</td>
</tr>
</tbody>
</table>

Note: Can ask the Index Advisor for suggestions...!
Let’s Jump to the Couchbase Slides