

# CS300 Couchbase NoSQL Server Administration

July 26th , 2016



PARTICIPANT USE ONLY IN CLASS 07.2015

Copyright © 2014 Couchbase, Inc. All rights Reserved. NOT FOR DISTRIBUTION

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Labs – Day 3



## Lab #5 (Server Warmup, Views and Indexes):

- Viewing server warmup metrics with cbstats
- Locate the warmup access log file on disk
- Use cbepctl tool to change alog\_sleep\_time and flush\_param settings
- Introduction to views in the beer-sample bucket
- Learn how to filter results when running a view
- Write a custom View with Map and Reduce code
- See analytics and metrics for a view
- Learn how to compact index files for a view
- Locate view/index files on disk
- Query a view and change its settings for how often the indexer runs via REST API

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Labs – Day 3



## Lab #6 (XDCR):

- Create a remote 2-node Couchbase cluster
- Configure un/encrypted, unidirectional/bi-directional replication between the 4-node & 2-node cluster
- Learn that design docs and views do not get replicated
- Examine XDCR conflict resolution and revision counts
- Discuss tombstones in the data files
- Differentiate between XDCR version 1 and version 2 protocols
- Examine advanced XDCR settings
- View outbound and inbound replication statistics
- View internal XDCR settings via REST API
- Examine about optimistic threshold setting for replication streams
- Delete replication streams

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).



# Warmup & Access Scanner

PARTICIPANT USE ONLY IN CLASS 02.2015

Copyright © 2014 Couchbase, Inc. All rights Reserved. NOT FOR DISTRIBUTION

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Server Warmup



- When a node is restarted, server must first warmup before it can handle client requests for data. If a client tries to contact Couchbase Server during warmup, the server will produce a **ENGINE\_TMPFAIL (0x0d)** error code.
- During warmup the server loads data from disk into RAM
- Couchbase can be configured to begin serving data before it has actually loaded all the keys and data from vBuckets
- An Access Log helps prioritize which frequently-used items to load into RAM first and is created by a process that will periodically scan every key in RAM and compile them into a log, named access.log as well as maintain a backup of this access log, named access.old
- By default this Scanner Process runs initially at 2:00 GMT and will run again in 24- hour time periods after that

4

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com)

This Engine\_Fail error indicates that data access is still not available because warmup has not yet finished. For those of you who are creating your own Couchbase SDK, you will need to handle this error in your library. This may mean that the client waits and retries, or the client performs a backoff of requests, or it produces an error and does not retry the request.

The server pre-fetches a list of most-frequently accessed keys and fetches these documents before it fetches any other items from disk.

The server can use this backup file during warmup if the most recent access log has been corrupted during warmup or node failure.

## Scanner Process



- Determines which keys are most frequently-used
- You can change the initial time from 2:00 GMT and/or the interval for the process
- Useful if you have a peak time for your application when you want the keys used during this time to be quickly available after server restart
- You can use cbepctl flush\_param to change the initial time and the interval for the process.

5

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

If you want to change this setting for an entire Couchbase cluster, you will need to perform this command on per-node and per-bucket in the cluster. By default, any setting you change with cbepctl will only be for the named bucket at the specific node you provide in the command.

## Scanner Process



To run the process every 20 mins or to change initial time:

```
# /opt/couchbase/bin/cbepctl localhost:11210 -b  
beer_sample set flush_param alog_sleep_time 20  
  
# /opt/couchbase/bin/cbepctl hostname:11210 -b beer_sample  
-p beer_password set flush_param alog_task_time 13
```

6

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

This is an opportunity to use a broad cluster command such as cluster shell.

<http://cea-hpc.github.io/clustershell/>

In this example we set the initial time to 1:00 PM UTC.

## Warmup



Couchbase Server is able to serve information from RAM when one of the following conditions is met during warmup:

- The server has finished loading documents for all keys listed in the access log, or
- The server has finished loading documents for every key stored on disk for all vBuckets, or
- The percentage of documents loaded into memory is greater than, or equal to, the setting for `ep_warmup_min_items_threshold`, or
- If total % of RAM filled by documents is greater than, or equal to, the setting for `ep_warmup_min_memory_threshold`, or
- If total RAM usage by a node is greater than or equal to the setting for `mem_low_watermark`,

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

When the server reaches one of these states, this is known as the *run level*; when Couchbase Server reaches this point, it immediately stops loading documents for the remaining keys. After this point, Couchbase Server will load this remaining documents from disk into RAM as a background data fetch.

## Warmup – cbstats command



```
# cbstats localhost:11210 -b beer_sample \
    -p bucket_password all | grep warmup

# cbstats hostname:11210 -b my_bucket \
    -p bucket_password raw warmup
```

**ep\_warmup\_thread** - Indicates whether the warmup completed or is still running. Returns “running” or “complete”.

8

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

You can use cbstats to get information about server warmup, including the status of warmup and whether warmup is enabled.

## Warmup Phases 4.x



### Initialization

Create Vbuckets for each Shard

Estimate Database ItemCount

Key Dump

Check for Access Log

Load Access Log/Load KV Pairs Data

LoadData

Done

9

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## 4.0 Warmup

The following stages in Warmup have been parallelized at a shard level:

- KeyDump
- LoadingKVPairs
- LoadingAccessLog
- LoadingData

Therefore at each of these stages, there will a **shard** number of process doing warmup on the vbuckets that correspond to each shard rather than the single process over the entire vbucket map. The end of each of these phases is marked by the completion of all the threads that were spawned at the level. Here's a sample state transition diagram that tells how warmup works now.

## ep\_warmup\_min\_items\_threshold



This indicates the percentage of items loaded in RAM that must be reached for Couchbase Server to begin serving data.

The lower this number, the sooner your server can begin serving data.

If you set this value to be too low, once requests come in for items, the item may not be in memory and Couchbase Server will experience cache-miss errors.

10

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).



# GSI Indexes and Views created using N1QL

PARTICIPANT USE ONLY IN CLASS 02/2015

Copyright © 2014 Couchbase, Inc. All rights Reserved. NOT FOR DISTRIBUTION

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Indexes Primary and Secondary



- A primary index is the ordering of the primary keys using UTF-8 encoding
- A secondary index is the ordering of some attribute(s) within a document(UTF-8)
- When created they are assigned to the least populated index nodes by default and can be specifically assigned to an index node when created
- The index tab will display indexes and command syntax Data Definition Language (DDL) used to create them

Such as:

- N1QL> CREATE PRIMARY INDEX ON `beer-sample` using gsi;
- or
- N1QL> create index Beer\_style on `beer-sample`(style) using gsi;

12

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

```
CREATE PRIMARY INDEX ON `travel-sample` using gsi;  
CREATE PRIMARY INDEX ON `gamesim-sample` using gsi;  
CREATE PRIMARY INDEX ON `beer-sample` using gsi;  
  
create index Beer_by_phone on `beer-sample`(phone) using gsi;  
create index Beer_style on `beer-sample`(style) using gsi;  
create index Beer_abv on `beer-sample`(abv) using gsi;  
create index Beer_by_state on `beer-sample`(state) using gsi;
```

Views can be created at the command line in N1QL as well using the “using view;” qualifier

These may have limited use with new index service.

```
CREATE PRIMARY INDEX ON `travel-sample` using view;  
CREATE PRIMARY INDEX ON `gamesim-sample` using view;  
CREATE PRIMARY INDEX ON `beer-sample` using view;  
create index Beer_by_phone on `beer-sample`(phone) using view;  
create index Beer_style on `beer-sample`(style) using view;  
create index Beer_abv on `beer-sample`(abv) using view;  
create index Beer_by_state on `beer-sample`(state) using view;
```



## Global service Indexes (GSI) and Views Using N1QL

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

```
create index Email on `Customer`(`Email) using gsi;
```

Create index on 'Email' field for documents in 'Customer' bucket  
Statement returns after index is built

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Now let's take a look at create index DDL statement, Data Definition Language (DDL).  
This statement will create the index definition as well as building the index  
Once the index is build, this statement will return

## Index DDL



```
create index Email1 on `Customers`(Email1) with  
{"defer_build":true} using gsi;
```

```
create index Phone1 on `Customers`(Phone1) with  
{"defer_build":true} using gsi;
```

build index on `Customers`(Email1, Phone1) using gsi;

Use defer\_build option to build index in background

Can build multiple indexes simultaneously

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Sometimes, you want to build the index at later time.

You can use the defer\_build option.

You will need to use the build index statement to start building the index

Using this statement, you can build multiple indexes together.



Create index Composite on 'Customer'(State,Zip)  
using gsi;

drop index Customers.Email using gsi;

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

These are other examples of index DDL.

You can create primary index, composite index or drop index.

## N1QL command line prompt



On a query service node use the following

```
# /opt/couchbase/bin/cbq
```

Or from another node that has Couchbase loaded ,  
you can point to the query engine

```
cbq -engine=http://172.23.107.18:8093
```

Be aware that your path will need to be modified and  
that there another cbq command in the linux OS  
(/usr/sbin/cbq)

17

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Usage of /opt/couchbase/bin/cbq:

-engine="http://localhost:8093/": URL to the query service(cbq-engine). By default, cbq connects to: http://localhost:8093

Examples:

```
cbq
```

Connects to local query node. Same as: cbq -engine=http://localhost:8093

```
cbq -engine=http://172.23.107.18:8093
```

Connects to query node at 172.23.107.18 Port 8093

```
cbq -engine=https://my.secure.node.com:8093
```

Connects to query node at my.secure.node.com:8093 using secure https protocol.

-quiet=false: Enable/Disable startup connection message for the shell

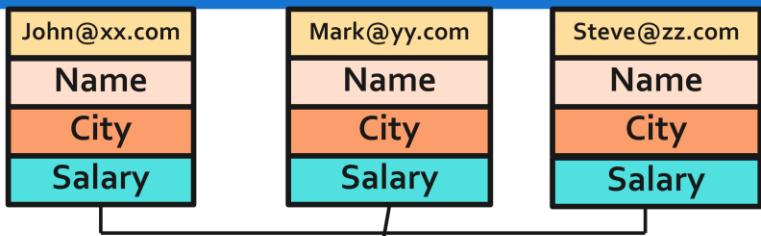
Default : false

Possible Values : true/false

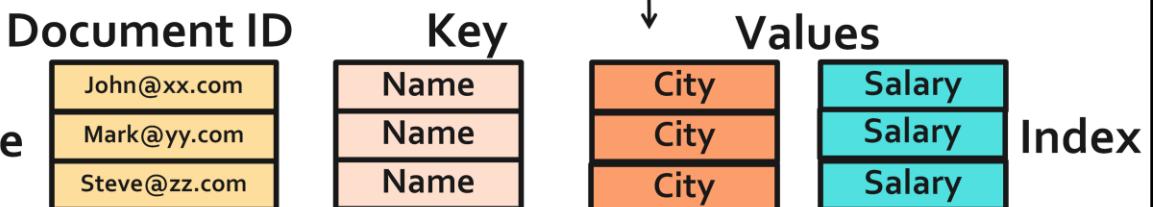
# Views and Indexes



## Documents



## Table



18

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

In the above example, the view takes the Name, City and Salary fields from the stored documents and then creates a array of this information for each document in the view. A view is created by iterating over every single document within the Couchbase bucket and outputting the specified information. The resulting index is stored for future use and updated with new data stored when the view is accessed. The process is incremental and therefore has a low ongoing impact on performance. Creating a new view on an existing large dataset may take a long time to build, but updates to the data will be quick.

# Views and Indexes



The screenshot shows the Couchbase Server interface with the 'beer-sample' database selected. The 'Views' tab is active. There are three design documents listed: '\_design/dev\_beer2' (language: javascript), 'brewery\_beers', 'by\_location', and '[Spatial] points'. For each document, there are buttons for 'Compact', 'Delete', 'Add Spatial View', 'Add View', 'Publish', 'Edit', and 'Delete'.

19

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

The final section provides a list of the design documents, and within each document, each defined view.

When viewing Development Views, you can perform the following actions:

Compact the view index with an associated design document. This will compact the view index and recover space used to store the view index on disk.

Delete a design document. This will delete all of the views defined within the design document.

Add Spatial View creates a new spatial view within the corresponding design document. See [Creating and Editing Views](#).

Add View creates a new view within the corresponding design document. See [Creating and Editing Views](#).

Publish your design document (and all of the defined views) as a production design document. See [Publishing Views](#).

# Views and Indexes



The screenshot shows the Couchbase Views and Indexes interface. At the top, there's a navigation bar with 'beer-sample' selected, followed by 'Views > \_design/dev\_beer2/\_view/brewery\_beers'. Below this, a document named '110F7F3614' is displayed with its JSON content. To the right of the document, a preview of the generated JSON output is shown. Below the document, there's a 'VIEW CODE' section containing Map and Reduce functions. The Map function is as follows:

```
Map
1 function(doc, meta) {
2   switch(doc.type) {
3     case "beer":
4       emit([meta.id]);
5       break;
6     case "brewery":
7       emit([doc.brewery_id, meta.id]);
8       break;
9   }
10 }
```

The Reduce function is listed as 'Reduce (built in: \_count, \_sum, \_stats)' with a value of 1.

At the bottom, there are 'Filter Results' options, a 'Show Results' button, and tabs for 'Development Time Subset' and 'Full Cluster Data Set'. A note says 'To see the results of this view, click "Show Results" above.'

20

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

The top portion of the interface provides navigation between the available design documents and views.

The Sample Document section allows you to view a random document from the database to help you write your view functions and so that you can compare the document content with the generated view output. Clicking the Preview a Random Document will randomly select a document from the database. Clicking Edit Document will take you to the Views editor

Document metadata is displayed in a separate box on the right hand side of the associated document. This shows the metadata for the displayed document, as supplied to the map() as the second argument to the function.

# Views and Indexes



The screenshot shows the Couchbase Views and Indexes interface. At the top, there are dropdown menus for 'beer-sample' and 'Views > \_design/dev\_beer2/\_view/brewery\_beers'. Below this is a search bar with the ID '110F222958' and buttons for 'Preview a Random Document' and 'Edit Document'. A 'VIEW CODE' button is also present. On the right side of the header are 'Save As...' and 'Save' buttons. Below the header, there is a 'Filter Results' section with a dropdown menu set to '?connection\_timeout=60000&limit=10&skip=0'. There are two tabs: 'Development Time Subset' (selected) and 'Full Cluster Data Set'. The main area displays a table with two columns: 'Key' and 'Value'. The keys listed are: ["110f0013c9"], ["110f0013c9", "110fd305e"], ["110f0013c9", "110fd3d0b"], ["110f0013c9", "110fd4296"], ["110f0013c9", "110fd4d92"], ["110f0013c9", "110fdde81"], ["110f0013c9", "110fd5443"], ["110f0013c9", "110fd56ff"], ["110fe0aaa7"], and ["110f001bbe"]. The corresponding 'Value' column for all rows is 'null'.

21

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

The top portion of the interface provides navigation between the available design documents and views.

The Sample Document section allows you to view a random document from the database so that you can compare the document content with the generated view output. Clicking the Preview a Random Document will randomly select a document from the database. If you know the ID of a document that you want to examine, enter the document ID in the box, and click the Lookup Id button to load the specified document.

To examine the function that generate the view information, use the View Code section of the display. This will show the configured map and reduce functions.

The lower portion of the window will show you the list of documents generated by the view. You can use the Show Results to execute the view.

The Filter Results interface allows you to query and filter the view results by selecting the sort order, key range, or document range, and view result limits and offsets.

# Views and Indexes



?connection\_timeout=60000&limit=10&skip=0

descending	<input type="checkbox"/>
startkey	<input type="text"/>
endkey	<input type="text"/>
startkey_docid	<input type="text"/>
endkey_docid	<input type="text"/>
group	<input type="checkbox"/>
group_level	<input type="text"/>
include_docs	<input type="checkbox"/>
inclusive_end	<input type="checkbox"/>
key	<input type="text"/>
keys	<input type="text"/>
reduce	<input type="radio"/> true <input type="radio"/> false <input type="radio"/> none
update_seq	<input type="checkbox"/>
stale	<input type="radio"/> false <input type="radio"/> update_after <input type="radio"/> ok
connection_timeout	<input type="text" value="60000"/>

22

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Clicking on the Filter Results query string will open a new window containing the raw, JSON formatted, version of the View results. To access the view results using the REST API.

# Views and Indexes



Views can be used within Couchbase Server for a number of reasons, including:

- Indexing and querying data from your stored objects
- Producing lists of data on specific object types
- Producing tables and lists of information based on your stored data
- Extracting or filtering information from the database
- Calculating, summarizing or reducing the information on a collection of stored data

23

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Since views can hold attribute(s) and will take up space on disk they should be used with on disk resource in mind.



## Global Secondary Indexes ( GSI )

### Created using N1QL

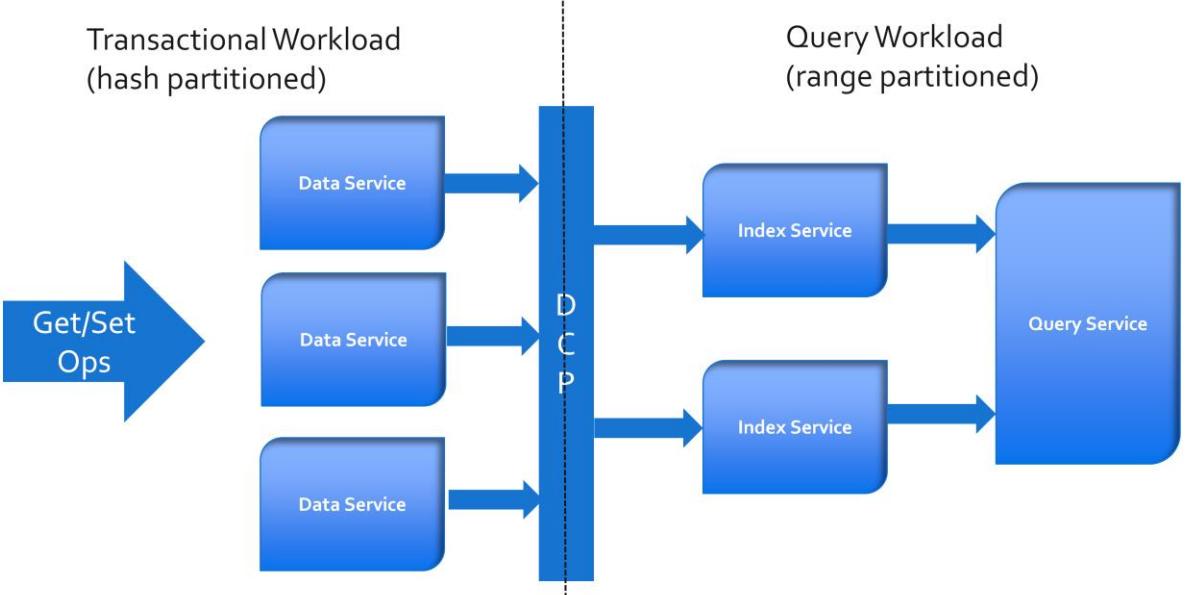
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Why GSI - Workload Isolation



Transactional Workload  
(hash partitioned)

Query Workload  
(range partitioned)



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

With Global secondary indexes(GSI)

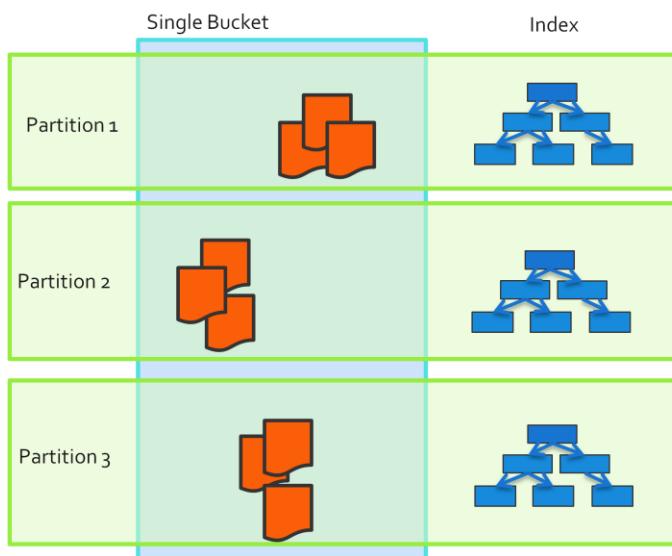
, you can separate query workload from transaction workload.

Using DCP, mutations in transactional workload is replicated to the index services.

And with this, you can add indexes to the query workload with minimal impact to your KV performance.

Indexes can be searched in memory to provide a much faster response than on disk

# Partitioned Index



```
create index IndexName on  
'BucketName'(FieldName)  
where FieldName="Value"  
using gsi;
```

Partitioned Index: Index on a subset of JSON documents

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

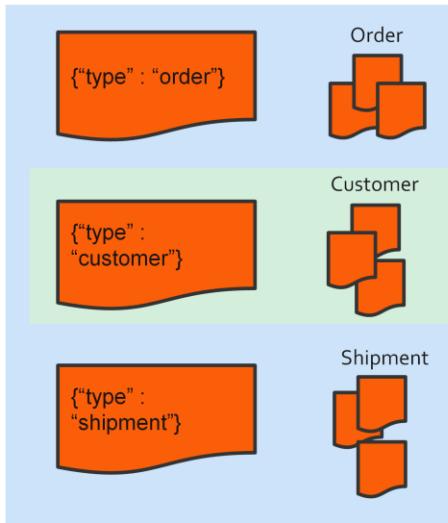
The next topic is partitioned index

- Definition
- One or more index per partition
- Where clause - expression

# Partitioned Index – Document Type



Bucket 'default'



```
create index Email on 'bucketName'(Email)  
where type="customer" using gsi;
```



Partitioned by document type (a field in document)  
Similar to index for a table

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

There are many ways to group documents.

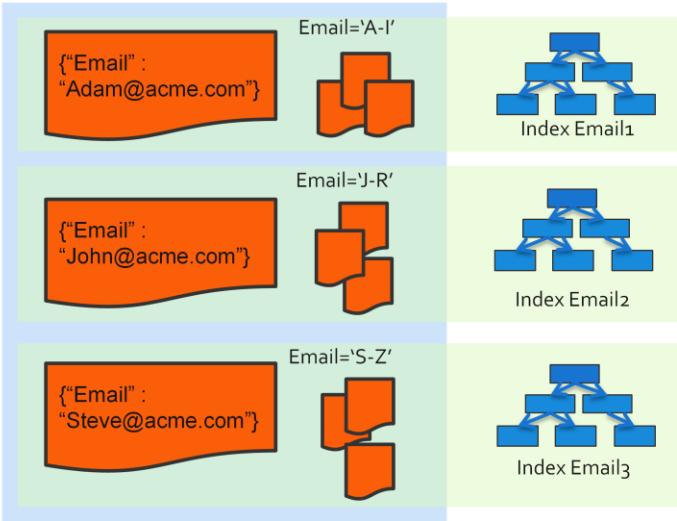
Here is one example of partitioned index where you partition by document type.

<example>

# Partitioned Index – Range



Bucket 'default'



```
create index Email1 on
'BucketName'(Email) where
Email > 'A' and Email <'I' using
gsi;
```

```
create index Email2 on
'BucketName'(Email) where
Email > 'J' and Email <'R' using
gsi;
```

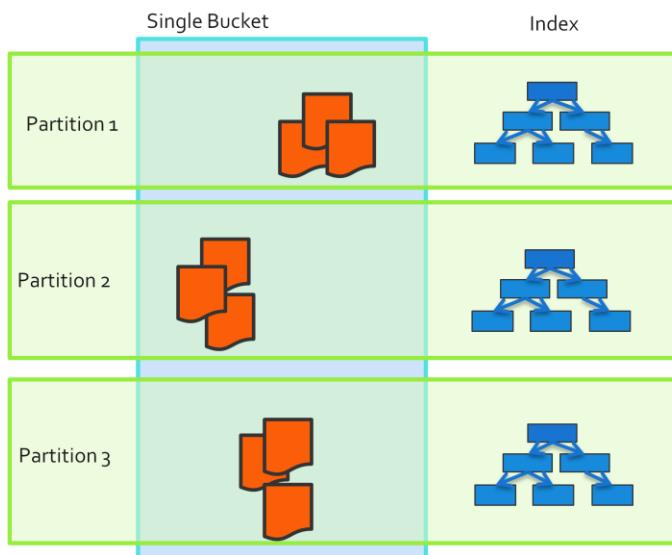
```
create index Email3 on
'BucketName'(Email) where
Email > 'S' and Email <'Z' using
gsi;
```

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Another example is to partition by range.

- Build index Email
- Break up in 3 ranges

# Partitioned Index - Benefits

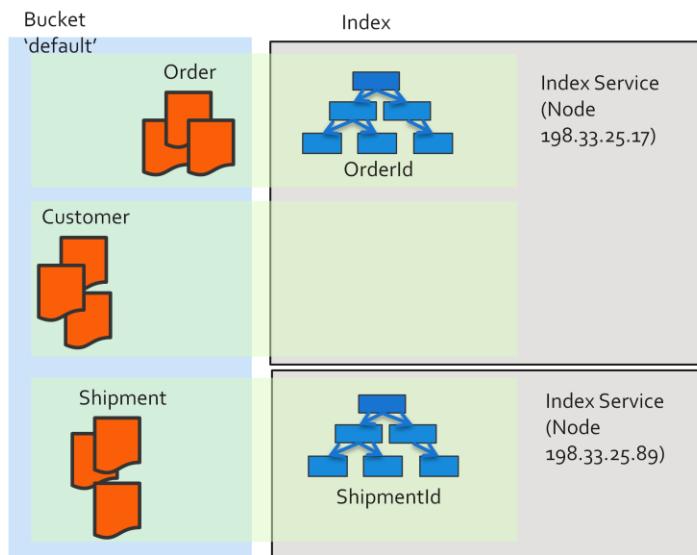


```
create index IndexName on  
'BucketName'(FieldName) where  
FieldName="Value" using gsi;
```

**Faster Index Build Time**  
**Lower Maintenance Latency**  
**Faster Scan Latency**  
**Higher Scan Throughput**

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Index Deployment - Default



```
create index OrderId on
  'BucketName'(OrderId) where
  type="order" using gsi;
```

```
create index ShipmentId on
  'BucketName'(ShipmentId) where
  type="shipment" using gsi;
```

**Default : on least populated node**

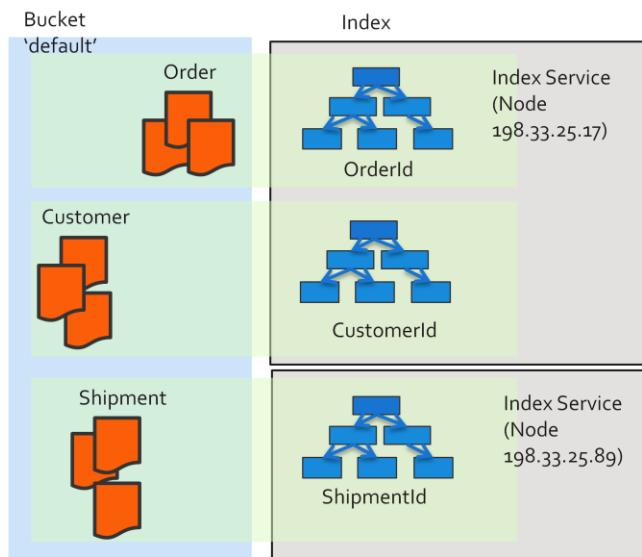
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

When you create an index, the indexer can decide for you where to put the index in the cluster

- Diagram
- Place index in the node with the less load
- Explicit control
- Index node IP
- With clause to specify the IP of index node



# Index Deployment - Explicit



```
create index CustomerId on
`BucketName`(CustomerId) where
type="order" with
{"nodes":"198.33.25.17"} using gsi;
```

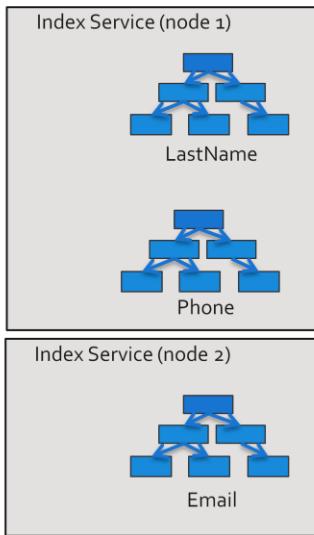
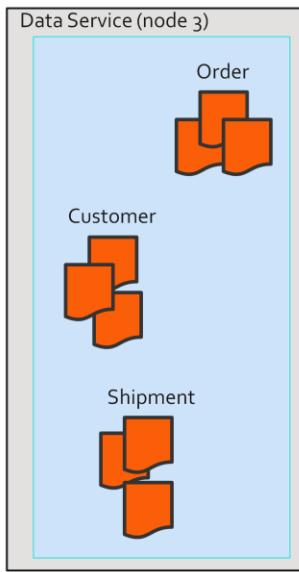
**Explicit : Specify node to deploy**

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

When you create an index, the indexer can decide for you where to put the index in the cluster

- Diagram
- Place index in the node with the less load
- Explicit control
- Index node IP
- With clause to specify the IP of index node

# Index Build



```
create index LastName on 'BucketName`('LastName) where type="customer" with {"defer_build": true} using gsi;
```

```
create index Phone on 'BucketName`('Phone) where type="customer" with {"defer_build": true} using gsi;
```

```
create index Email on 'BucketName`('Email) where type="customer" with {"defer_build": true} using gsi;
```

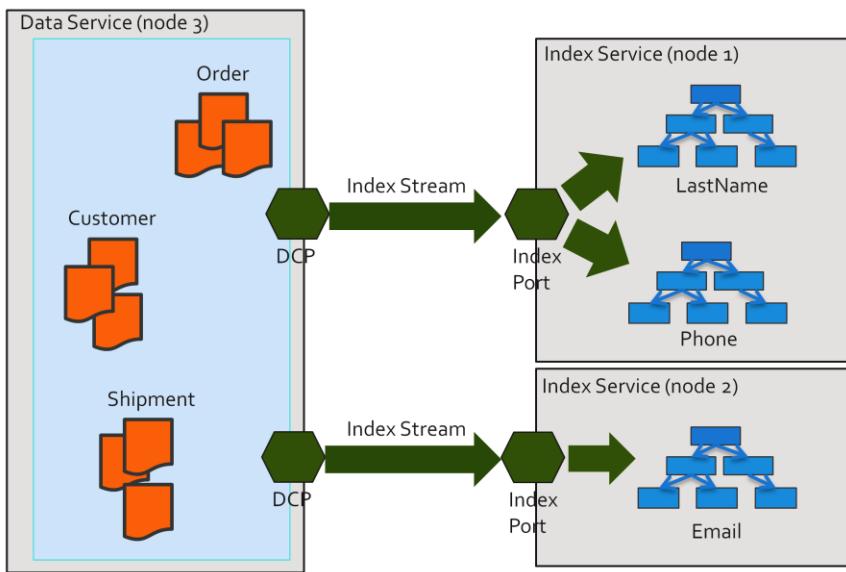
```
build index on 'BucketName`('LastName, Phone, Email) using gsi;
```

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

After creating index, let see how an index is built.

- Diagram
- 3 create index
- With defer option
- Build index to build the 3 indexes together

# Index Build – DCP



build index on  
'BucketName'  
'(LastName, Phone,  
Email) using gsi;

A dedicated DCP  
stream for each index  
service

Stream is shared among  
indexes within the  
same service

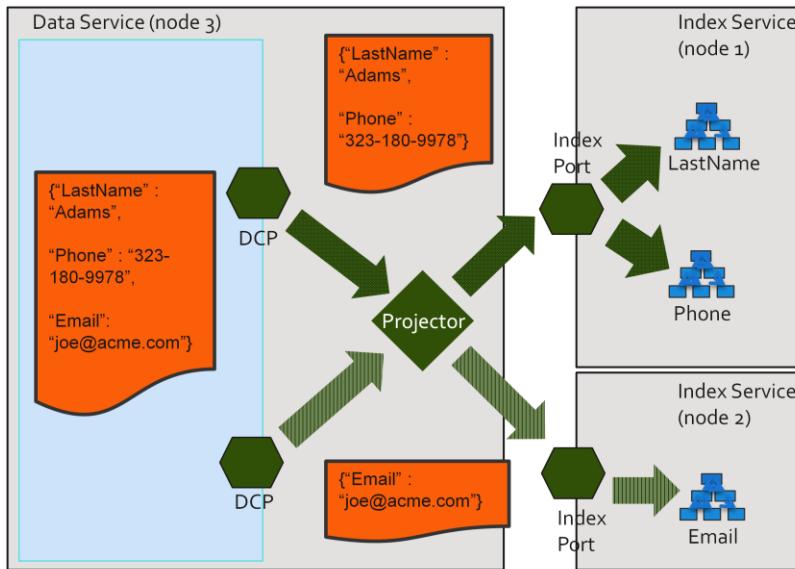
DCP streams are  
concurrent across index  
services

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

When building a new index, we will create a new DCP stream between each data service and index service.

- Indexes in the same node share the same DCP stream
- Last name and phone sharing the stream
- If you have two index services, you will have two concurrent DCP streams
- All the indexes can be built parallel.

# Index Build – Projector



**Projector resides in data service listening to DCP streams**

Extract secondary keys from documents based on index definition

Route secondary keys to appropriate index service based on index deployment

**Reduce Network Traffic**

Projector Payload = only keys relevant for a index service

Filter out irrelevant mutation

Payload size can be much smaller than document size

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Now let's look inside data service, we actually have a single broker that manage both DCP streams.

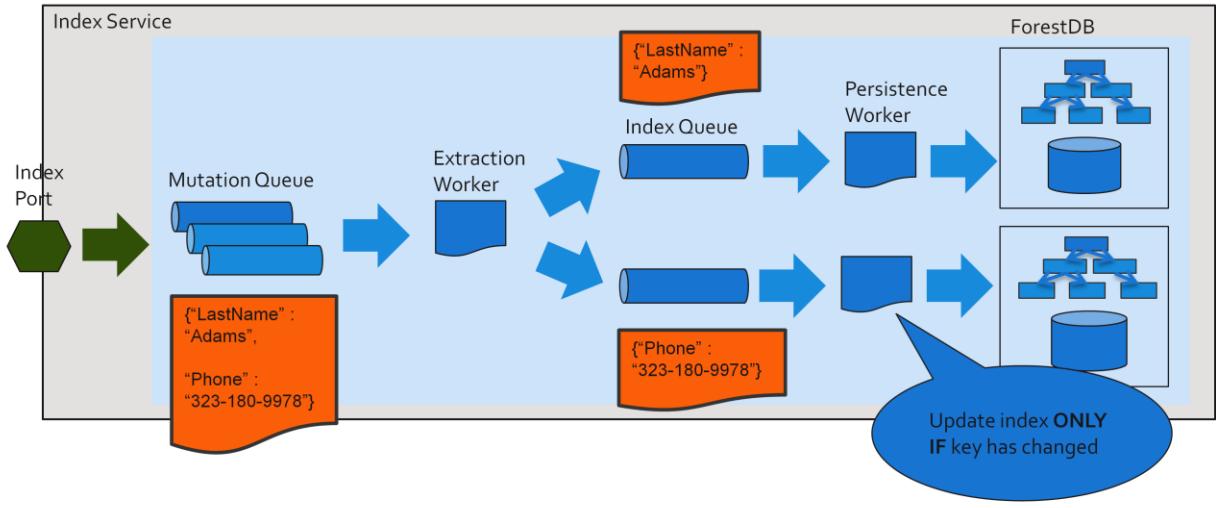
This broker is called projector.

The projector takes the mutation from DCP stream, extract the keys from the mutations, and route the keys to the appropriate index service.

The main purpose of the projector is to reduce network traffic by filtering out irrelevant mutations as well as irrelevant keys.

<example>

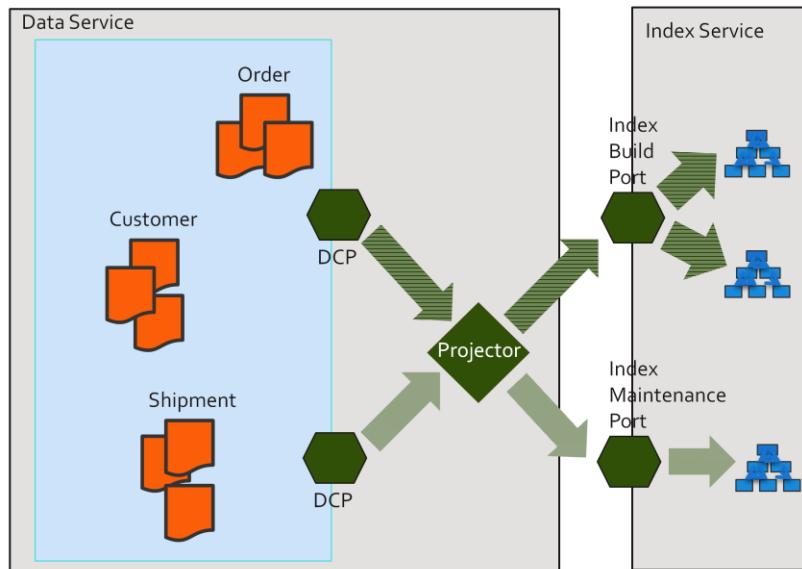
# Index Build – Parallel Indexer Pipeline



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

- Diagram
- This indexer pipeline is highly parallel.

# Index Maintenance



**Updating index with on-going mutations**

**Parallel DCP streams**

Concurrent index build and index maintenance

**One shared maintenance stream**

for all active indexes in each service

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Once the index is built, the index will be kept up-to-date with on-going mutations.  
It works the same way as index build.

The indexer will use a separate DCP stream for index maintenance for each index service.

All active indexes will be updated using this single DPC stream for the same service.  
Since it is using a different DCP stream, index build and index maintenance can run in parallel.



## Persisted Snapshot

For Rollback during Recovery  
(checkpointing)

Tunable snapshot interval

## In-memory Snapshot

- For scanning
- Query stability and consistency
- Tunable snapshot interval

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

During index build and index maintenance, the indexer will create snapshots.

A snapshot represents a version of an index.

There are 2 types of snapshots:

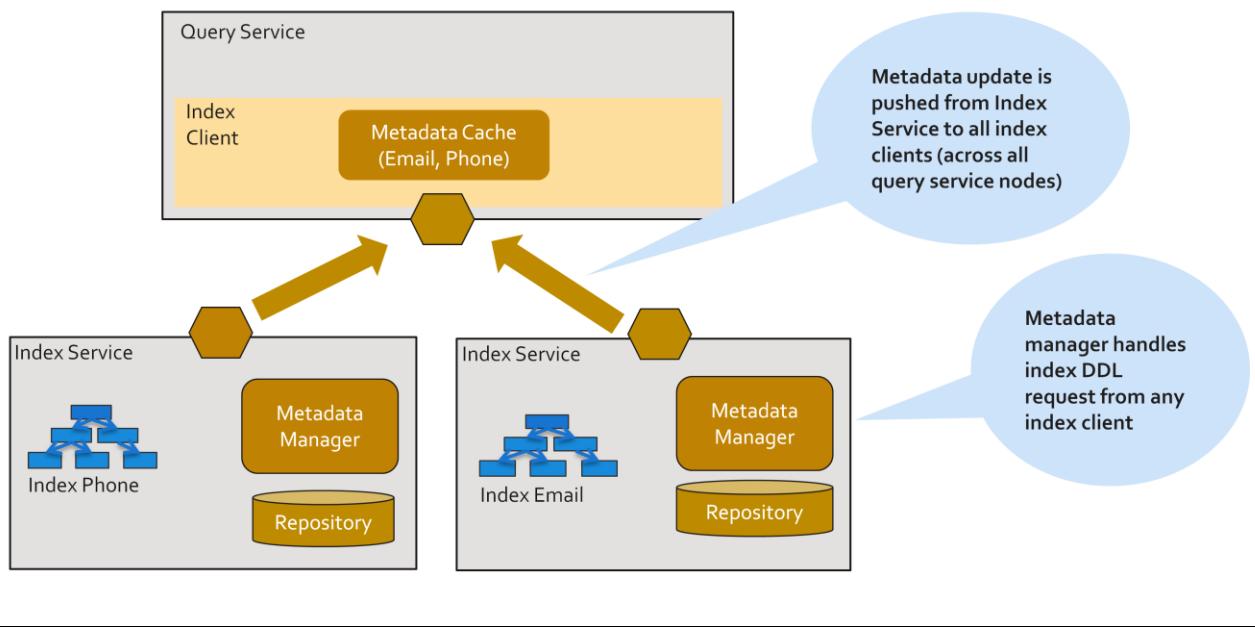
A persisted snapshot creates a version of the persisted data.

It is used for rollback during recovery.

A in-memory snapshot creates a version of the in-memory index data.

It is used to provide query stability and consistency.

# Index Metadata



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Inside each index service, there is a single metadata manager that handles index DDL for that service

This, in turn, will update the metadata repository in the index service.

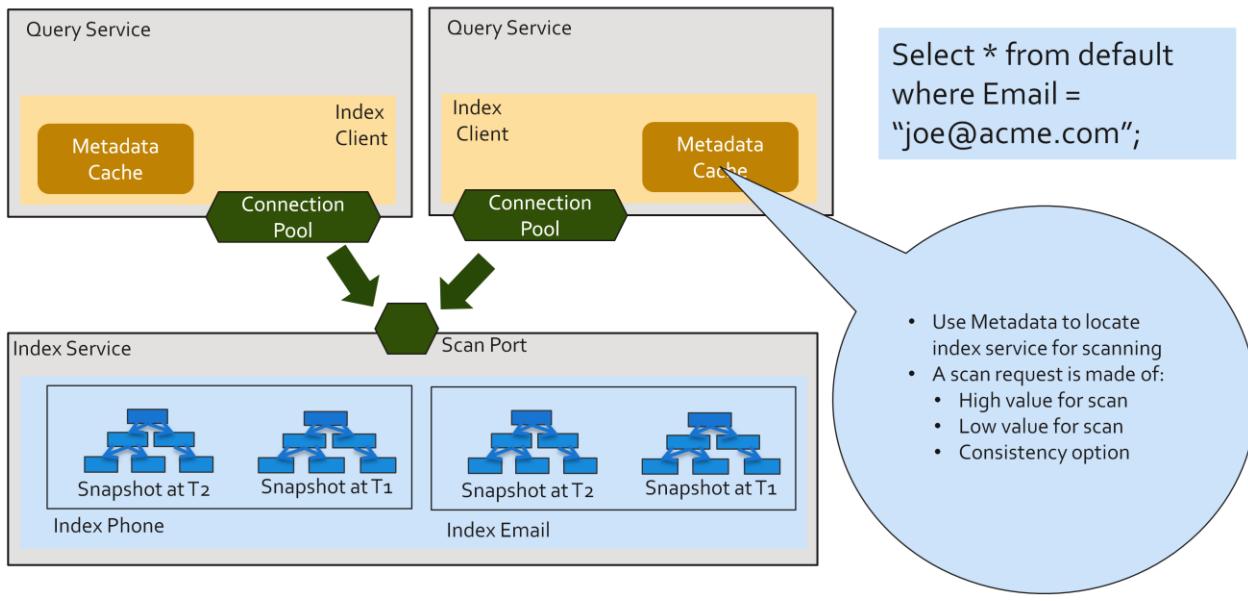
Once the index metadata is updated, it will be pushed to the query service.

Each query service has an index client.

The index client itself has a metadata cache which keeps track of all the index metadata.



# Index Scan



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

After an index is built, it will be in active state.

And you can now use that index for query.

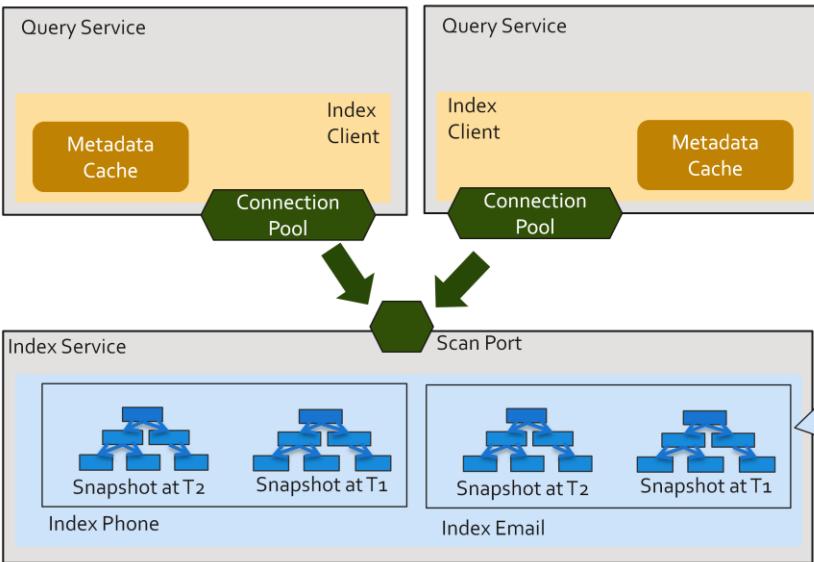
<explain the diagram>

When the query service wants to perform a scan, it will use the metadata to locate the physical location of the index service.

A request is then sent to the indexer. The request includes ...



# Index Scan



Select \* from default  
where Email =  
"joe@acme.com";

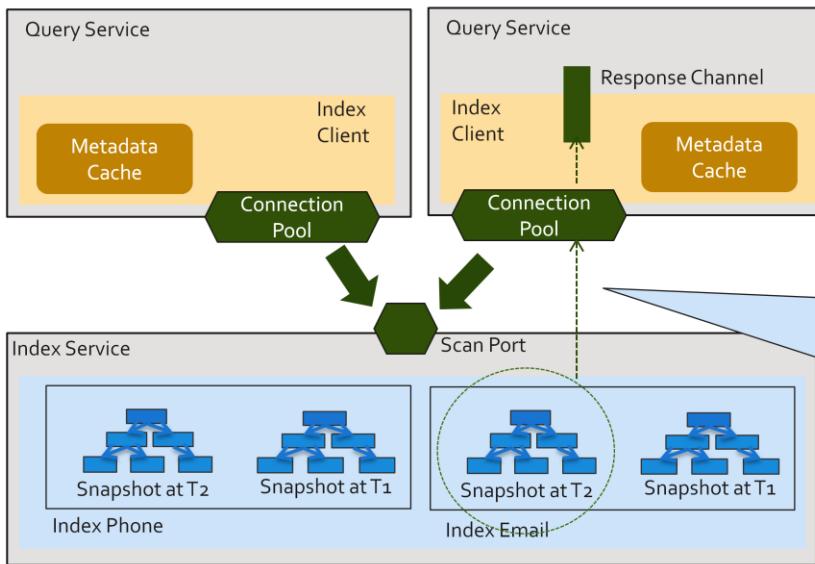
- Pick snapshot based on consistency option:
- Unbounded : Pick the latest snapshot in indexer
  - Request\_plus : Observe current or future snapshot that matches query request

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Once the request arrives at the index service, the indexer will use the consistency option to pick a snapshot for scan.

The 2 consistency options are ...

# Index Scan



Select \* from default  
where Email =  
"joe@acme.com";

- Open an iterator on the chosen snapshot with given range (min/max values).
- Stream result back to Index Client
- Results are placed in response channel

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).



## Covering indexes

42

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Covering indexes



When an index includes the actual values of all the fields specified in the query, the index covers the query and does not require an additional step to fetch the actual values from the data service.

An index, in this case, is called a covering index and the query is called a covered query. As a result, covered queries are faster and deliver better performance.

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## What are covering indexes in Couchbase?

Couchbase Server 4.1 brings covering indexes to N1QL. If you have an index that covers all of the data required by your N1QL query, Couchbase Server can return the data straight from that index. This cuts at least one, and possibly several, document reads from the query process.

How does it work?

When you submit a N1QL query to Couchbase, it goes through the following steps:

- Parses your query, analyses what it does and then creates a query plan.
- Checks the relevant indexes.
- Gets the keys for the documents that contain the results and then fetches the documents themselves.
- Pulls the relevant data out of those documents to create the result.
- Returns the result to your application.

With a covering index, all the data for the result is in the index itself so you cut out steps 3 and 4.

How do I use covering indexes?

Nothing changes about how you write your queries. Instead, this is all about how you set-up the indexes. A single index -- GSI or view -- must cover every JSON key that your query touches.

Got an example?

Let's say we're storing customer profiles. They might look like this:

```
{  
  "name": "Sherlock Holmes",  
  "address": "221b Baker Street",  
  "city": "London",  
  "country": "UK",  
  "newsletter": true,  
  "joinedDate": 1168041600  
}
```

If we wanted to find the name and city of everyone who is subscribed to our newsletter, we could write a N1QL query like this:

```
SELECT name, city FROM `customers` WHERE newsletter=true;
```

Without a covering index, this query would pull name and city from our customer records by doing a look-up of those documents.

We can create a covering index for this query like this:

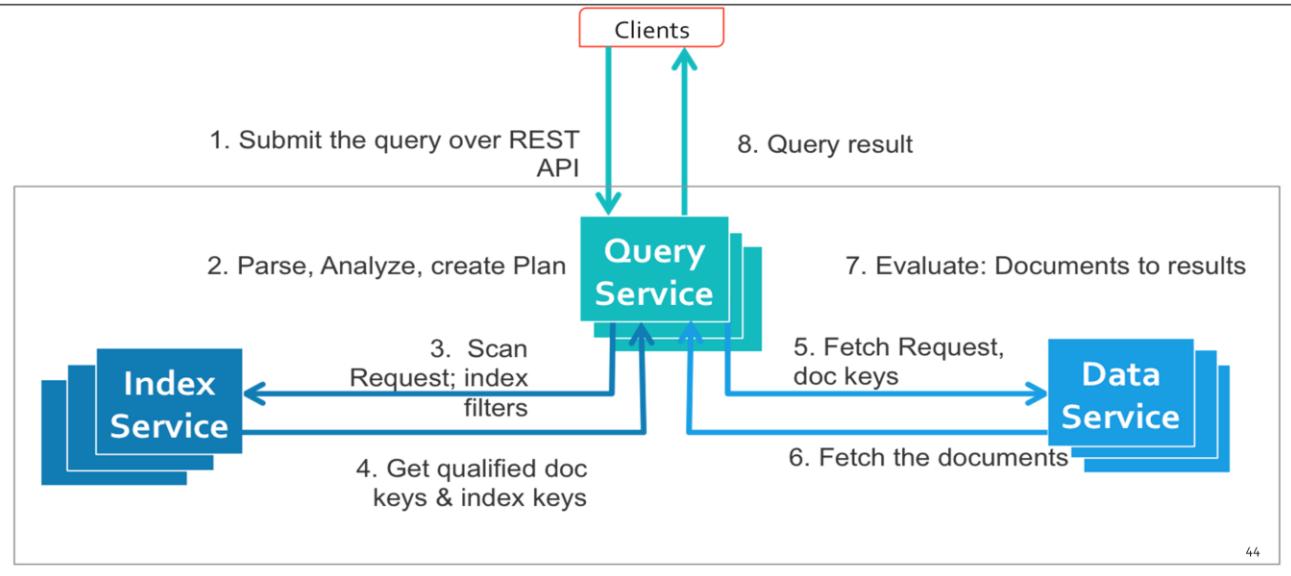
```
CREATE INDEX newsletter_subs ON `customers`(name, city, newsletter) USING GSI;
```

Now the query can be satisfied entirely from the index.

# Standard N1QL query



diagram illustrates query execution work flow w/o covering indexes:



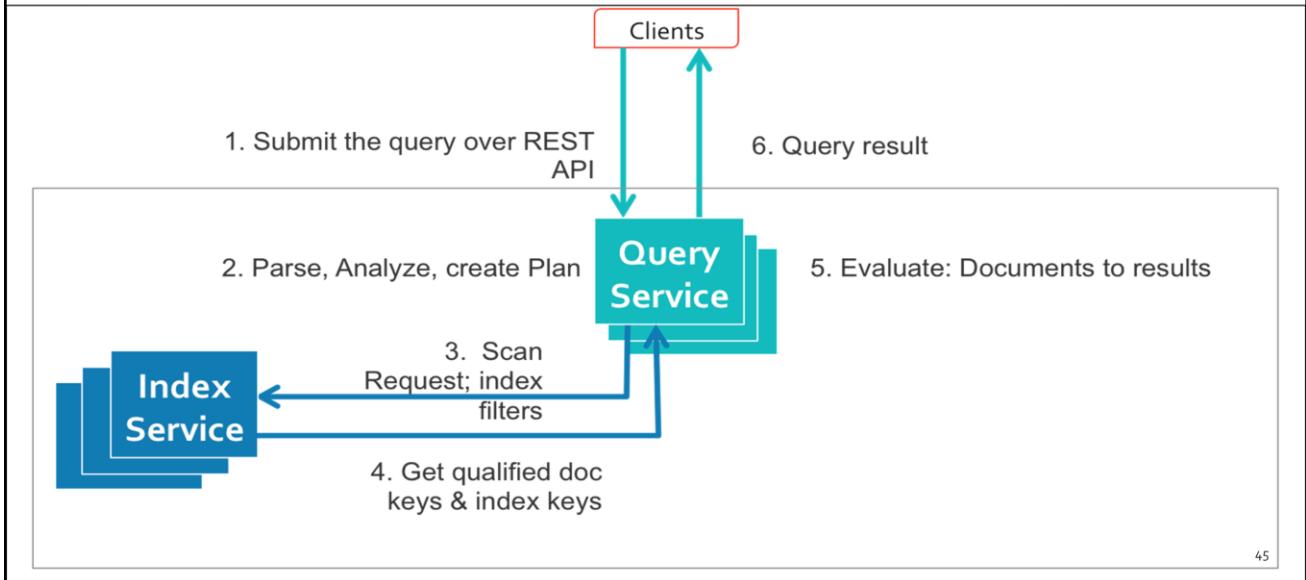
44

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Query with covering indexes



Diagram illustrates query execution work flow with covering indexes:



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

45



## Views

# Secondary Indexes Using Views (View tab)

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Views and Indexes



- A view creates an index on the stored information according to the format and structure defined within the view.
- The view consists of specific fields and information extracted from the objects stored in Couchbase.
- Views create indexes on your information allowing you to search and select information stored within Couchbase Server.
- The process is incremental and therefore has a low ongoing impact on performance. Creating a new view on an existing large dataset may take a long time to build, but updates to the data will be quick.
- Because the process relies on the fields of stored JSON, if the document is not JSON, or the requested field in the view does not exist, the information is ignored.

47

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

The purpose of a view is to take the un-structured, or semi-structured, data stored within your Couchbase Server database, extract the fields and information that you want, and to produce an index of the selected information.

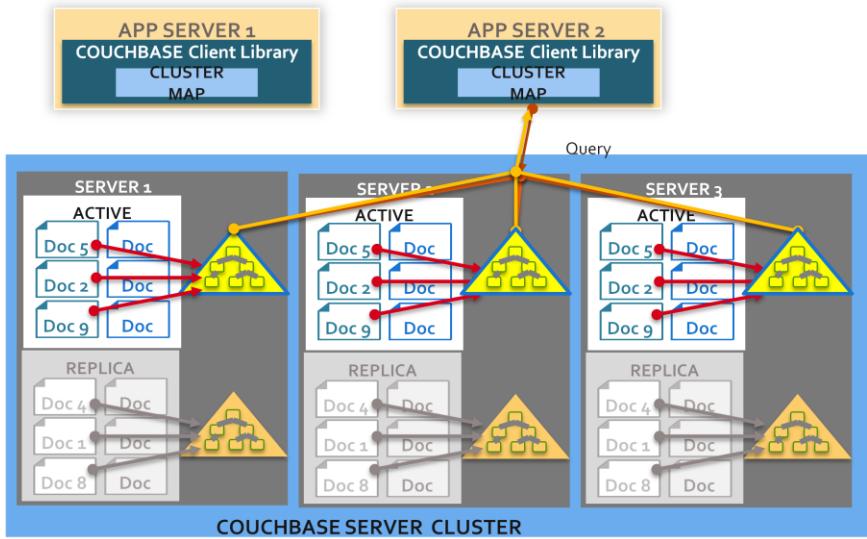
Views live on data service nodes on disk

Indexes live on Index nodes in memory and on disk for persistence

Views can be created Using MapReduce code on the view tab OR by using the N1QL command line.

Accessing views created by MapReduce using a different API than accessing views created using N1QL

# Client Interaction During Querying of a view



- Indexing work is distributed amongst nodes
- Single node accepts request
- Request distributed to other nodes
- Queries combine the results from required nodes
- Data available in client as hash/array (implementation dependent)

User Configured Replica Count = 1

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

On Disk Views use the scatter gather method with any node able field a query but spread the logic out to all nodes and gathering results.

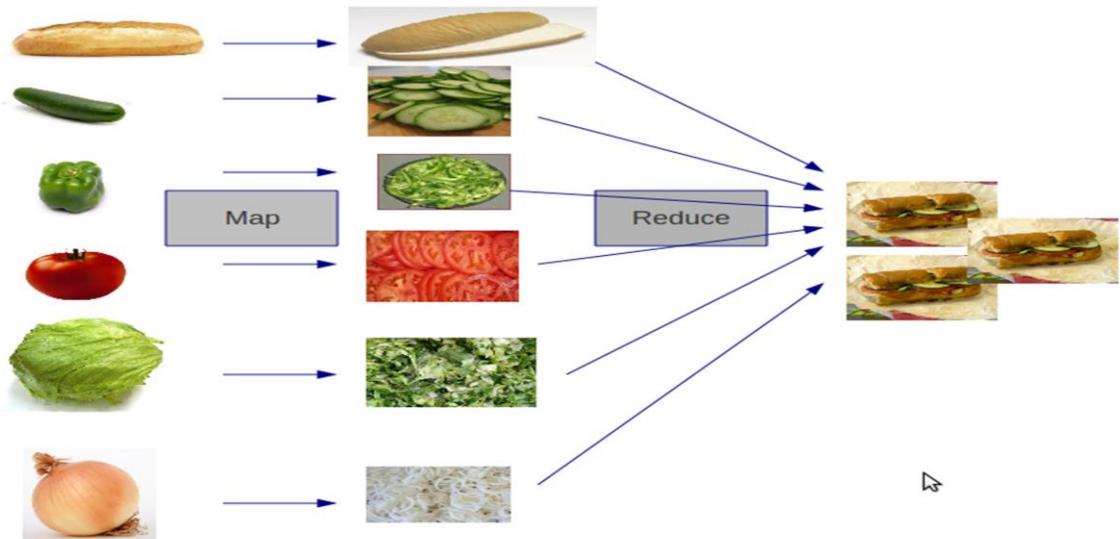
# Indexing Subsystems



- **Storing data and Indexing data are separate systems in all databases**
- **In explicit schema scenarios (RDBMS), Indexes are optimized based on the data type(s)**
- **In flexible schema scenarios Map-Reduce is used to create views**

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# What is Map-Reduce?



50

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# What is Map-Reduce ?



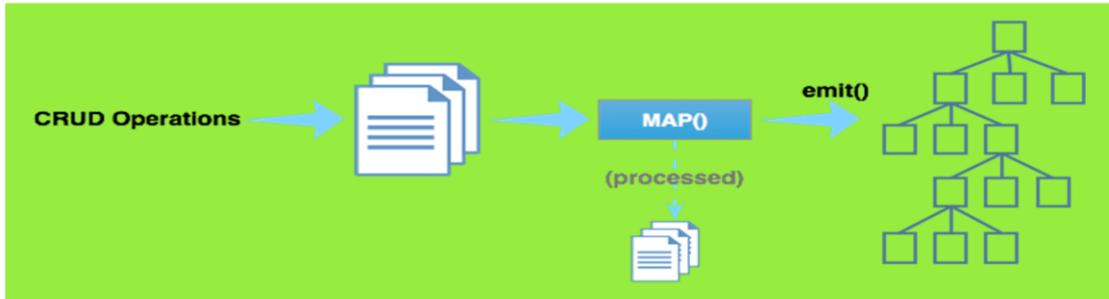
- **Map-Reduce is a technique designed for dealing with Big Data and processing in parallel in distributed systems**
- **Map-Reduce is also specifically designed for dealing with unstructured or semi-structured data**
- **Map functions identify data with collections, process them, and output transformed values**
- **Reduce functions take the output of Map functions and perform numeric aggregate calculations on them**

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Views: Map-Reduce Indexes



- In Couchbase, Map-Reduce is specifically used to create Indexes
- Map functions are applied to JSON documents and they output or "emit" data that is organized in an Index form



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## Sample View



```
function (doc, meta) {  
    if (doc.type == "beer" && doc.brewery_id && doc.name) {  
        emit(doc.name, doc.abv);  
    }  
}
```

- Creates an Index of Beer Names (doc.name) and the Alcohol By Volume values (doc.abv)
  - Filters Documents
    - Only JSON Documents with json key doc.type == "beer"
    - and doc.brewery\_id is non-null
    - and doc.name is non-null
  - Outputs
    - Beer Name (doc.name) [searchable]
    - Beer Alcohol By Volume (doc.abv) [row value]

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Why Use Map-Reduce Indexes?



- Index (Find) Documents by different JSON Values
- Query Documents by JSON Values
- Create Statistics and Aggregates

## When are Indexes Necessary?

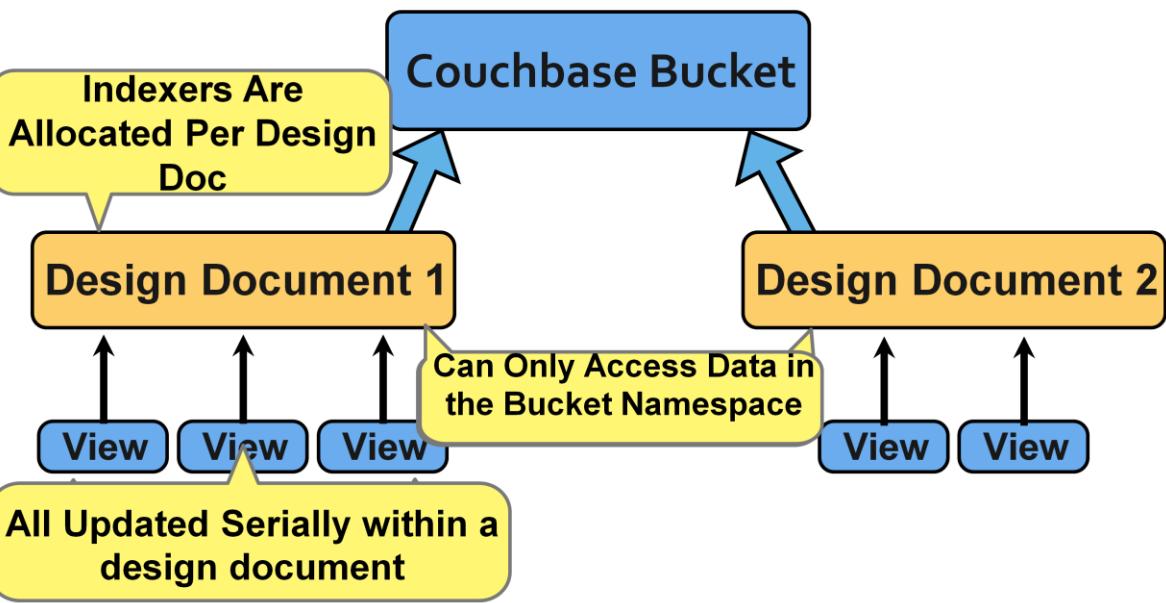
- Documents are Keyed by Random Properties (UUID, GUID, etc.)
- Iterating through Lists of Documents with Random Keys
- Iterating through Lists of Documents on different JSON Properties (i.e. all User docs, all Product docs, by Timestamp, etc.)

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Documents that are stored with an expiry are not automatically removed until the background expiry process removes them from the database.

This means that expired documents may still exist within the index.

## Buckets >> Design Documents >> Views



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

All the views within a single design document are updated when the update to a single view is triggered. For example, a design document with three views will update all three views simultaneously when just one of these views is updated.

## Updates:



Updates can be triggered in two ways:

- At the point of access or query by using the stale parameter
- Automatically by Couchbase Server based on the number of updated documents, or the period since the last update.
- Automatic updates can be controlled either globally, or individually on each design document.

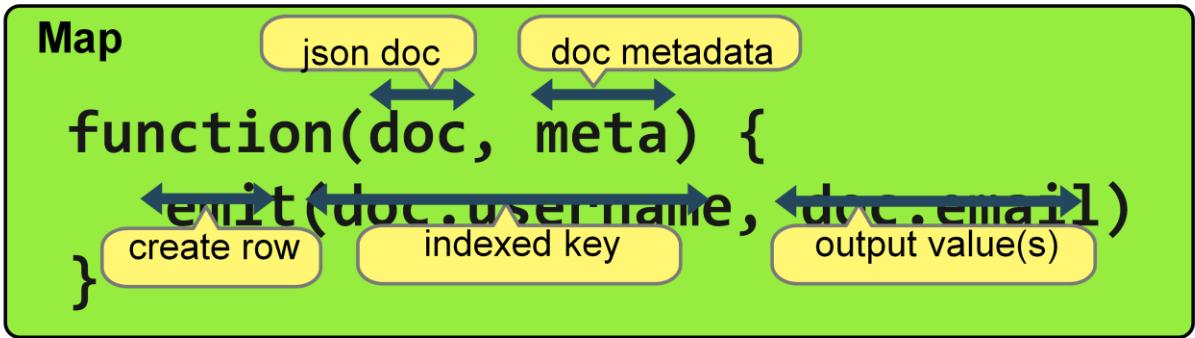
56

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Map() Function => Index



Every Document passes through View Map() functions



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Single Element Keys (Text Key)



## Map

```
function(doc, meta) {  
    emit<del>doc.email, doc.points  
}  
                                         ↑  
                                         text key
```

meta.id	doc.email	doc.points
u::1	abba@couchbase.com	1000
u::35	jasdeep@couchbase.com	1200
u::20	zorro@couchbase.com	900

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Compound Keys (Array)



Array Based Index Keys get sorted as Strings,  
but can be grouped by array elements

## Map

```
function(doc, meta) {  
    emit(dateToArray(doc.timestamp),  
        1)  
}
```

meta.id	dateToArray(doc.timestamp)	value
u::20	[2012,10,9,18,45]	1
u::1	[2012,9,26,11,15]	1
u::35	[2012,8,13,2,12]	1

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## Primary Index (as a view)

To create a ‘primary key’ index, i.e. an index that contains a list of every document within the database, with the document ID as the key, you can create a simple view:

```
Function (doc ,meta)
{
    emit (meta.id, null) ;
}
```

This enables you to iterate over the documents stored in the database.

This will provide you with a view that outputs the document ID of every document in the bucket using the document ID as the key.

```
?startkey="object100" &endkey="object199"
```

PARTICIPANT USE ONLY IN CLASS Q3 2015

Copyright © 2014 Couchbase, Inc. All rights Reserved. NOT FOR DISTRIBUTION

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).



60

To create a ‘primary key’ index, i.e. an index that contains a list of every document within the database, with the document ID as the key, you can create a simple view:

# Secondary Index



For example, given the document structure:

```
{  
  "firstname": "Martin",  
  "lastname": "Brown"  
}
```

A view to support queries on the `firstname` field could be defined as follows:

```
function(doc, meta)  
{  
  if (doc.firstname)  
  {  
    emit(doc.firstname.toLowerCase(),null);  
  }  
}
```

Queries can now be specified by supplying a string converted to lowercase. For example:

```
?key="martin"
```

Will return all documents where the `firstname` field contains 'Martin', regardless of the document field capitalization.

61

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

The view works as follows for each document:

- Only outputs a record if the document contains a `firstname` field.
- Converts the content of the `firstname` field to lowercase.

# Using expiration metadata



You can use the expiration field within a view to get the list of recently active sessions in a system:

```
function(doc, meta)
{
  if (doc.type && doc.type == "session")
  {
    emit(meta.expiration, doc.nickname)
  }
}
```

62

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

The metadata object makes it very easy to create and update different views on your data using information outside of the main document data. For example, you can use the expiration field within a view to get the list of recently active sessions in a system.

If you have sessions which are saved with a TTL, this will allow you to give a view of who was recently active on the service.



# Querying Views and Indexes

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# View Query Parameters



- **key = ""**
  - used for exact match of index-key
- **keys = []**
  - used for matching set of index-keys
- **startkey/endkey = ""**
  - used for range queries on index-keys
- **startkey\_docID/endkey\_docID = ""**
  - used for range queries on meta.id
- **stale=[false, update\_after, true]**
  - used to decide indexer behavior from client
- **group/group\_by**
  - used with reduces to aggregate with grouping

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Most Common Query's Are Ranges



?startkey="bz" & endkey="zn"

Pulls the Index-Keys  
between UTF-8 Range  
specified by the  
startkey and endkey.



doc.email	meta.id
abba@couchbase.com	u::1
beta@couchbase.com	u::7
jasdeep@couchbase.com	u::2
math@couchbase.com	u::5
matt@couchbase.com	u::6
yeti@couchbase.com	u::4
zorro@couchbase.com	u::3

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Index-Key Matching



?key="math@couchbase.com"

"  
Match a Single Index-  
Key



doc.email	meta.id
abba@couchbase.com	u::1
beta@couchbase.com	u::7
jasdeep@couchbase.com	u::2
math@couchbase.com	u::5
matt@couchbase.com	u::6
yeti@couchbase.com	u::4
zorro@couchbase.com	u::3

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Index-Key Set Matches



?keys=[“math@couchbase.com”,  
“yeti@couchbase.com”]  
Query Multiple in the  
Set (Array Notation)



doc.email	meta.id
abba@couchbase.com	u::1
beta@couchbase.com	u::7
jasdeep@couchbase.com	u::2
math@couchbase.com	u::5
matt@couchbase.com	u::6
yeti@couchbase.com	u::4
zorro@couchbase.com	u::3

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Unicode ordering compared to Byte ordering



## Byte Order

1234567890 < a-z < A-Z

## Unicode Collation

1234567890 < aAbBcCdDeEfFgGhHiljJkKlLmM...  
a < á < A < Á < b

If it were Byte Order 2 Queries Merged:

`startkey="y"& endkey="z" merged with startkey="Y"&endkey="Z"`

With Unicode Collation gets both y and Y:

`startkey="y"&endkey="z"`

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Couchbase Server uses a Unicode collation algorithm to order letters, so you should be aware of how this functions. Most developers are typically used to Byte order, such as that found in ASCII and which is used in most programming languages for ordering strings during string comparisons.

The first line shows: The following shows the order of precedence used in Byte order, such as ASCII:

This means any items that start with integers will appear before any items with letters; any items that begin with capital letters will appear before items in lower case letters. This means the item named “Apple” will appear before “apple” and the item “Zebra” will appear before “apple”.

Compare this with the order of precedence used in Unicode collation, which is used in Couchbase Server: Notice again that items that start with integers will appear before any items with letters. However, in this case, the lowercase and then uppercase of the same letter are grouped together. This means that if “apple” will appear before “Apple” and would also appear before “Zebra”.

# Ordering continued.....



## Ordering Example

In Byte order, keys in an index would appear as follows:

```
"ABC123" < "ABC223" < "abc123" < "abc223" < "abcd23" < "bbc123" < "bbcd23"
```

The same items will be ordered this way by Couchbase Server under Unicode collation:

```
"abc123" < "ABC123" < "abc223" < "ABC223" < "abcd23" < "bbc123" < "bbcd23"
```

**This is particularly important for you to understand if you query Couchbase Server with a startkey and endkey to get back a range of results.**

69

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Fix the Byte ordered screenshot, it's incorrect.

# Understanding the Stale request flag



**stale = UPDATE\_AFTER (default if nothing is specified)**

always get fastest response

can take two queries to read your own writes

**stale = OK**

auto update will trigger eventually

might not see your own writes for a few minutes

least frequent updates -> least resource impact

**stale = FALSE**

Use with Persistence observe if data needs to be included in view results

BUT aware of delay it adds, only use when really required

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

When accessing a view, the contents of the view are asynchronous to the stored documents. In addition, the creation and updating of the view is subject to the stale parameter. This controls how and when the view is updated when the view content is queried.

`update_after : Allow stale view, update view after access`

`ok : Allow stale view`

`false : Force update of the view index before results are returned`

## Built-In Reduces



- Are faster than creating your own reduces for the same information
  - **\_count**
    - gives count for number of items in Index
  - **\_sum**
    - sums value parameters (for numeric values only)
  - **\_stats**
    - gives sum, count, min, max and sum of squares for statistics

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## Custom Reduces



- Are a bit tricky at first, it's a skill!
- Learn about it through our docs, practice first, most common problem in custom reduces is that they don't "reduce" the data
- Can be creatively used!
- Always do it in a separate Design Document to sandbox it from your existing Views, if you have a logic problem or error it won't interrupt existing Views

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).



## Sample Buckets

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Beer and Brewery Data



- demonstrates a combination of the document structure used to describe different items, including references between objects, and also includes a number of sample views that show the view structure and layout.

```
The primary document type is the 'beer' document:  
  
{  
    "name": "Piranha Pale Ale",  
    "abv": 5.7,  
    "ibu": 0,  
    "srn": 0,  
    "upc": 0,  
    "type": "beer",  
    "brewery_id": "110f04166d",  
    "updated": "2010-07-22 20:00:20",  
    "description": "",  
    "style": "American-Style Pale Ale",  
    "category": "North American Ale"  
}
```

74

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Beer documents contain core information about different beers, including the name, alcohol by volume ( abv ) and categorization data.

## Beer and Brewery Data



- Individual beer documents are related to brewery documents using the brewery\_id field, which holds the information about a specific brewery for the beer:

```
{  
    "name": "Commonwealth Brewing #1",  
    "city": "Boston",  
    "state": "Massachusetts",  
    "code": "",  
    "country": "United States",  
    "phone": "",  
    "website": "",  
    "type": "brewery",  
    "updated": "2010-07-22 20:00:20",  
    "description": "",  
    "address": [  
    ],  
    "geo": {  
        "accuracy": "APPROXIMATE",  
        "lat": 42.3584,  
        "lng": -71.0598  
    }  
}
```

75

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

The brewery record includes basic contact and address information for the brewery, and contains a spatial record consisting of the latitude and longitude of the brewery location.

## brewery\_beers view

- view outputs a composite list of breweries and beers they brew by using the view output format to create a 'fake' join
- This outputs the brewery ID for brewery document types, and the brewery ID and beer ID for beer document types:

```
function(doc, meta) {  
    switch(doc.type) {  
        case "brewery":  
            emit([meta.id]);  
            break;  
        case "beer":  
            if (doc.brewery_id) {  
                emit([doc.brewery_id, meta.id]);  
            }  
            break;  
    }  
}
```

```
{  
    "total_rows" : 7315,  
    "rows" : [  
        {  
            "value" : null,  
            "id" : "110f0013c9",  
            "key" : [  
                "110f0013c9"  
            ]  
        },  
        {  
            "value" : null,  
            "id" : "110fdd305e",  
            "key" : [  
                "110f0013c9",  
                "110fdd305e"  
            ]  
        },  
        {  
            "value" : null,  
            "id" : "110fdd3d0b",  
            "key" : [  
                "110f0013c9"  
            ]  
        }  
    ]  
}
```

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

76

The output could be combined with the corresponding brewery and beer data to provide a list of the beers at each brewery.



## by\_location view

- Outputs the brewery location, accounting for missing fields in the source data. The output creates information either by country, by country and state, or by country, state and city.

```
function (doc, meta) {
  if (doc.country, doc.state, doc.city) {
    emit([doc.country, doc.state, doc.city], 1);
  } else if (doc.country, doc.state) {
    emit([doc.country, doc.state], 1);
  } else if (doc.country) {
    emit([doc.country], 1);
  }
}
```

77

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

The view also includes the built-in \_count function for the reduce portion of the view

## by\_location view



- The view also includes the built-in \_count function for the reduce portion of the view. Without using the reduce, the information outputs the raw location information:

```
{  
    "total_rows" : 1413,  
    "rows" : [  
        {  
            "value" : 1,  
            "id" : "110f0b267e",  
            "key" : [  
                "Argentina",  
                "",  
                "Mendoza"  
            ]  
        },  
        {  
            "value" : 1,  
            "id" : "110f035200",  
            "key" : [  
                "Argentina",  
                "Buenos Aires",  
                "San Martin"  
            ]  
        }  
    ]  
}
```

78

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## by\_location view



- With the reduce() enabled, grouping can be used to report the number of breweries by the country, state, or city. For example, using a grouping level of two, the information outputs the country and state counts:

```
{"rows": [ {"key": ["Argentina", ""], "value": 1}, {"key": ["Argentina", "Buenos Aires"], "value": 1}, {"key": ["Aruba"], "value": 1}, {"key": ["Australia"], "value": 1}, {"key": ["Australia", "New South Wales"], "value": 4}, {"key": ["Australia", "NSW"], "value": 1}, {"key": ["Australia", "Queensland"], "value": 1}, {"key": ["Australia", "South Australia"], "value": 2}, {"key": ["Australia", "Victoria"], "value": 2}, {"key": ["Australia", "WA"], "value": 1} ] }
```

79

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

How many entries for New South Wales? 4 or 5 ?

With fields allowing multi string entries New South Wales fits but also allows a field entry of NSW(short for new south wales)

Application developer would need to control for this.

## Settings tab

## Sample Buckets (sub tab)



```
# ls /opt/couchbase/samples
beer-sample.zip gamesim-sample.zip travel-sample.zip
```

80

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

This is a repeated slide from Day 1

Sample buckets can be loaded with attendant indexes and views.

These can be deleted and reloaded later should the need for development arise.

If you are doing a rebalance you may wish to delete these to speed the rebalance and then load them later.

Directory for including your own sample databases is given here.



## BEER SAMPLE VIEW

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Beer Sample Database Example



meta

```
{  
  "id": "00000000-0000-0000-0000-000000000000",  
  "rev": 0,  
  "expiration": 0,  
  "flags": 0,  
  "type": "json"  
}
```

```
"name": "Aventine",  
"abv": 8.2,  
"ibu": 0,  
"m": 0,  
"type": "beer",  
"brewery_id": "110f1f2012",  
"updated": "2010-07-22 20:00:20",  
  
"description": "Dark-ruby, almost black-colored and streaked with fine top-  
fermenting yeast, this beer has a compact and persistent head. This is a very intense  
wheat doppelbock with a complex spicy chocolate-like aroma with a hint of banana and  
raisins. On the palate, you experience a soft touch and on the tongue it is very rich  
and complex, though fresh with a hint of caramel. It finishes in a rich soft and  
lightly bitter impression.",  
  
"style": "South German-Style Weizenbock",  
  
"category": "German Ale"
```

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Map Function - Index Definition



## ▼ VIEW CODE

### Map

```
1  function (doc, meta) {  
2    if (doc.type == "beer" && doc.brewery_id) {  
3      emit(doc.brewery_id, doc.abv);  
4    }  
5  }
```

The code snippet shows a Map function. It checks if the document type is 'beer' and has a brewery ID. If so, it emits the brewery ID as the indexed key and the document's ABV (Alcohol By Volume) as the value(s). The code is annotated with three yellow boxes: '+row' under the opening brace, 'indexed key' under the emit statement, and 'value(s)' under the doc.abv parameter. Double-headed green arrows connect '+row' to 'indexed key' and 'indexed key' to 'value(s)'.

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Result Set - Brewery ID's by Beer



Key	Value
"110f0013c9" 110fdd305e	brewery_id 5 . 2
"110f0013c9" 110fdd3d0b	document key (of the beer) 5 . 8
"110f0013c9" 110fdd4296	7
"110f0013c9" 110fdd4d92	6 . 8
"110f0013c9" 110fdd4e81	6
"110f0013c9" 110fdd5443	7 . 6

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Reduce Values (doc.abv) with \_stats



## ▼ VIEW CODE

### Map

```
1 function (doc, meta) {  
2   if (doc.type == "beer" && doc.brewery_id) {  
3     emit(doc.brewery_id, doc.abv);  
4   }  
5 }
```

### Reduce (built in: \_count, \_sum, \_stats)

```
1 _stats|
```

add \_stats built-in reduction

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Query with Group and Reduce



Find average alcohol by volume per brewery.

**VIEW CODE**

```
Map
1  function (doc, meta) {
2    if(doc.type === "beer" && doc.brewery_id) {
3      emit(doc.brewery_id, doc.abv);
4    }
5 }
```

**Reduce (built in: \_count, \_sum, \_stats)**

```
_stats
```

**Save As... Save**

**Filter Results**  **?group=true&reduce=true&stale=false&connection\_timeout=60000&limit=10&skip=0** **Show Results**

**Development Time Subset** **Full Cluster**

**set group=true & reduce=true**

Key	Value
"110f0013c9"	{ "sum": 54.5999999999999, "count": 8, "min": 5.2, "max": 8.2, "sumsqr": 380.92 }
"110f001bbe"	{ "sum": 63.7000000000001, "count": 11, "min": 3.6, "max": 9.8, "sumsqr": 393.8700000000001 }
"110f002955"	{ "sum": 11, "count": 2, "min": 5, "max": 6, "sumsqr": 61 }
"110f0032cc"	{ "sum": 22.8, "count": 5, "min": 0, "max": 5.9, "sumsqr": 130.12 }

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Groups Brewery\_ID's, Reduces for Stats



Filter Results

?group=true&reduce=true&connection\_timeout=60000&limit=10&skip=0

Development Time Subset

Full Cluster

group=true & reduce=true

Key

Value

"110f0013c9"  
undefined { "sum": 54.59999999999999, "count": 8, "min": 5.2, "max": 8.2, "sumsqr": 380.92 }

"110f001bbe"  
undefined { "sum": 63.7, "count": 11, "min": 3.6, "max": 9.8, "sumsqr": 393.87 }

"110f0032cc"  
undefined { "sum": 22.8, "count": 5, "min": 0, "max": 5.9, "sumsqr": 130.12 }

"110f004251"  
undefined { "sum": 13.4, "count": 2, "min": 6.6, "max": 6.8, "sumsqr": 89.7999999999998 }

"110f004c2a"  
undefined { "sum": 24, "count": 3, "min": 6, "max": 10, "sumsqr": 200 }

Brewery ID's are Grouped, and \_stats collected (Reduced)

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

If Sum = 11 and Count = 2, the Average is 5.5

## Lab #5 (Server Warmup, Views and Indexes)



- Viewing server warmup metrics with cbstats
- Locate the warmup access log file on disk
- Use cbepctl tool to change alog\_sleep\_time and flush\_param settings
- Introduction to views in the beer-sample bucket
- Learn how to filter results when running a view
- Write a custom View with Map and Reduce code
- See analytics and metrics for a view
- Run compaction on index files for a view
- Locate view/index files on disk
- Query a view and change its settings for how often the indexer runs via REST API



Time: 1 hour

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).



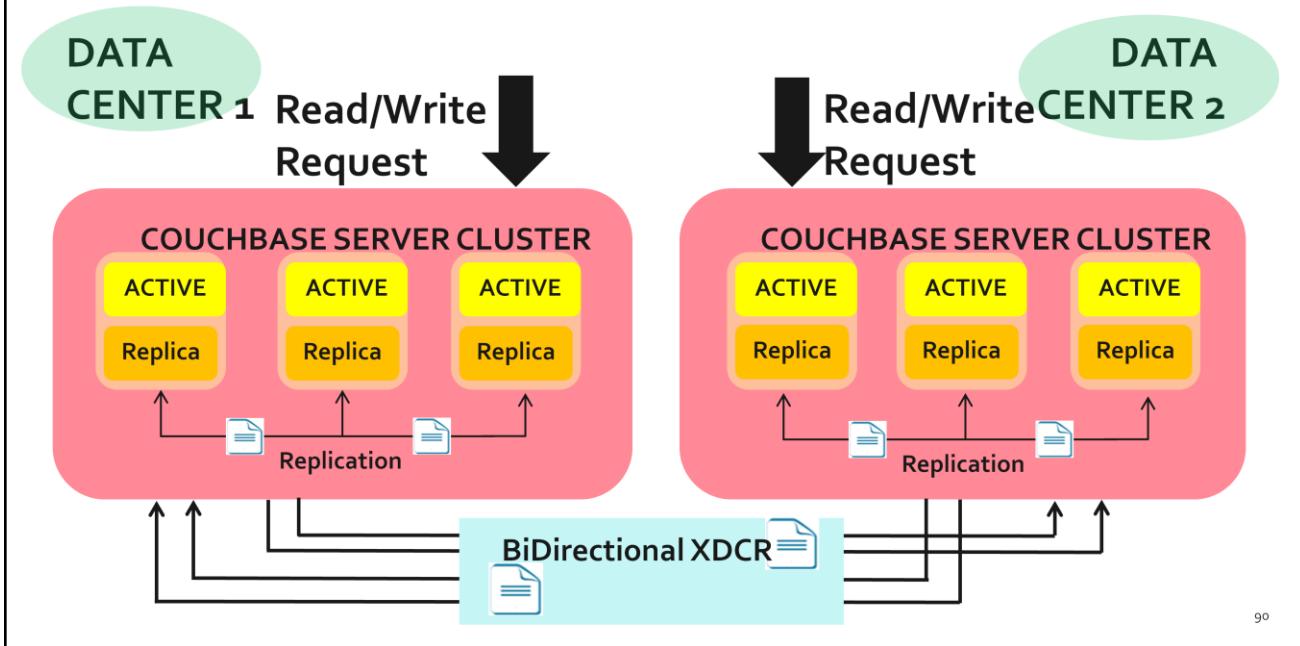
**XDCR**

**PARTICIPANT USE ONLY IN CLASS Q3 2015**

Copyright © 2014 Couchbase, Inc. All rights Reserved. **NOT FOR DISTRIBUTION**

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Cross datacenter replication in Couchbase Server



90

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

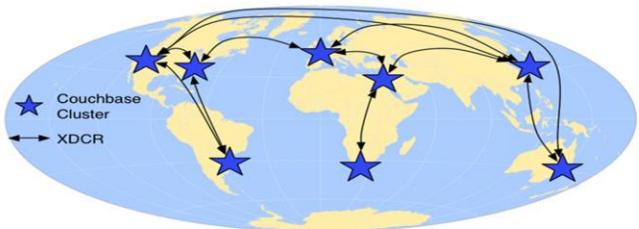
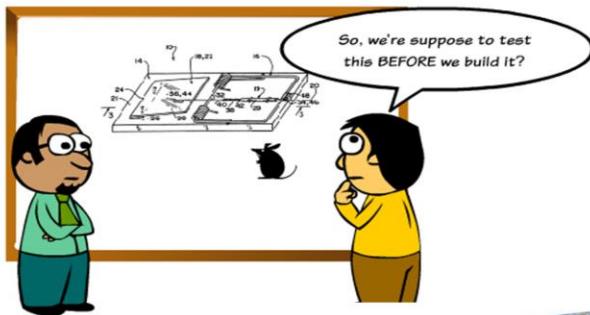
## XDCR Use Cases

Disaster Recovery.

Bringing Data Closer to Users.

Data Replication for Development and Test Needs.

# XDCR Use Cases



91

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

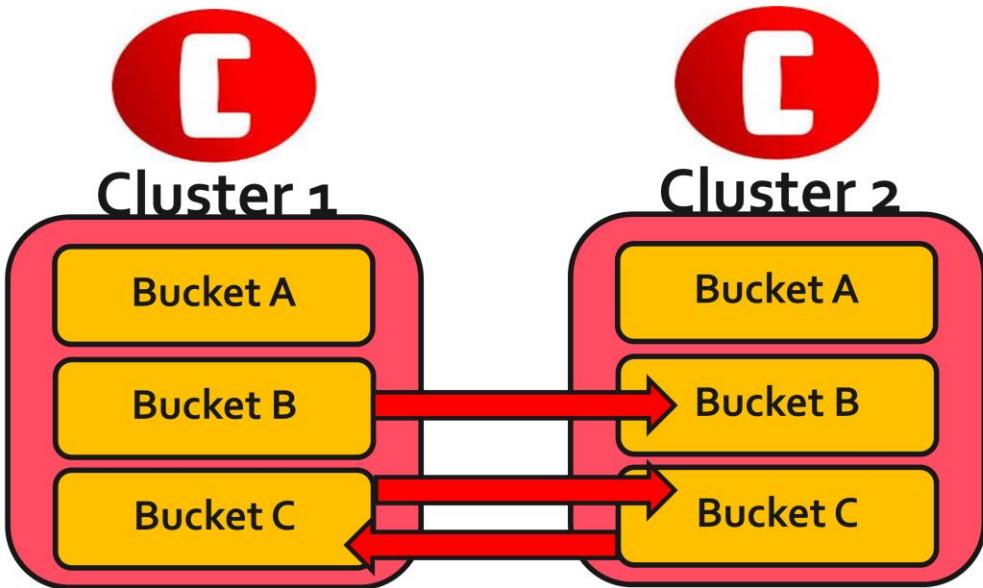
## XDCR Use Cases

### Disaster Recovery.

### Bringing Data Closer to Users.

### Data Replication for Development and Test Needs.

# Cross Datacenter Replication (XDCR)



92

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## Basic Topologies

XDCR can be configured to support a variety of different topologies; the most common are unidirectional and bidirectional. Unidirectional Replication is one-way replication, where active data gets replicated from the source cluster to the destination cluster. You may use unidirectional replication when you want to create an active offsite backup, replicating data from one cluster to a backup cluster. Bidirectional Replication allows two clusters to replicate data with each other. Setting up bidirectional replication in Couchbase Server involves setting up two unidirectional replication links from one cluster to the other. This is useful when you want to load balance your workload across two clusters where each cluster bidirectionally replicates data to the other cluster.

XDCR can be setup on a per bucket basis. A bucket is a logical container for documents in Couchbase Server. Depending on your application requirements, you might want to replicate only a subset of the data in Couchbase Server between two clusters. With XDCR you can selectively pick which buckets to replicate between two clusters in a unidirectional or bidirectional fashion.

# Replication Fundamentals



- **Unidirectional**
  - **Create Backup Cluster**
  - **Create Geographic Cluster, Chained Propagation**
  - **Create Development/Test Cluster with Active Data**
  - **Integrate with Elastic Search**
  - **Integrate with Custom Consumer**
- **Bi-Directional**
  - **Populate Multiple Clusters with Data Modified in each Cluster**

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

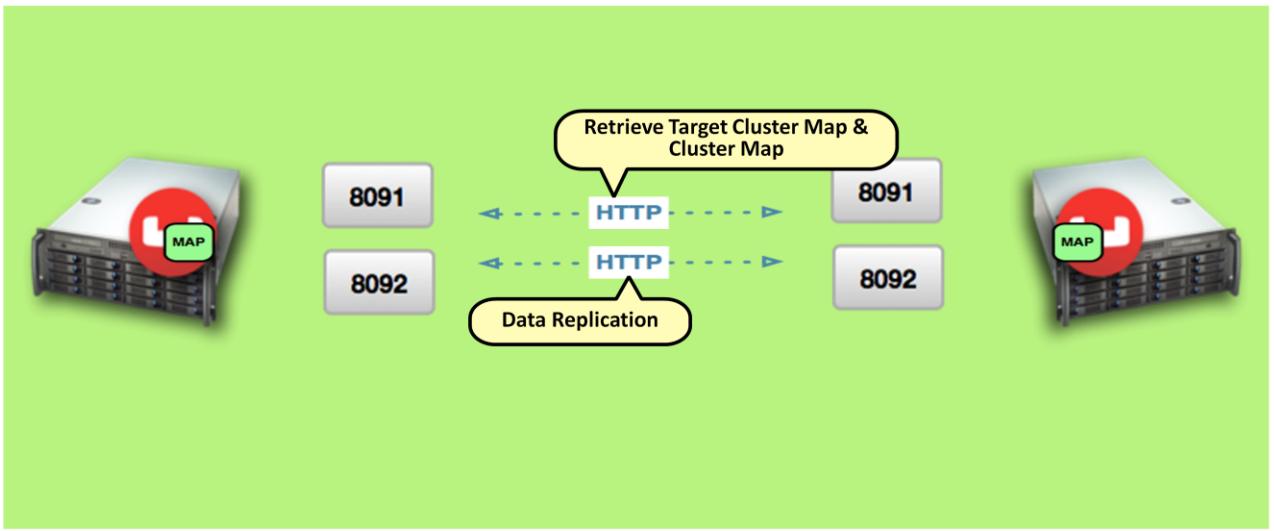
# Active-Active Replication



- XDCR configured Per Bucket
- Within Bucket, Each Partition (vbucket) Replicated Directly
  - Data is replicated from Active Partition -> Active Partition
  - Each Partition (1024) has a separate append only file
- XDCR process:
  - . Queues documents for replication
  - . Transmits Documents via HTTP in batches
  - . Keeps track of Checkpoint Intervals if Replication Streams are Disconnected
  - . Upon disconnection between clusters, automatically handles reconnection attempts

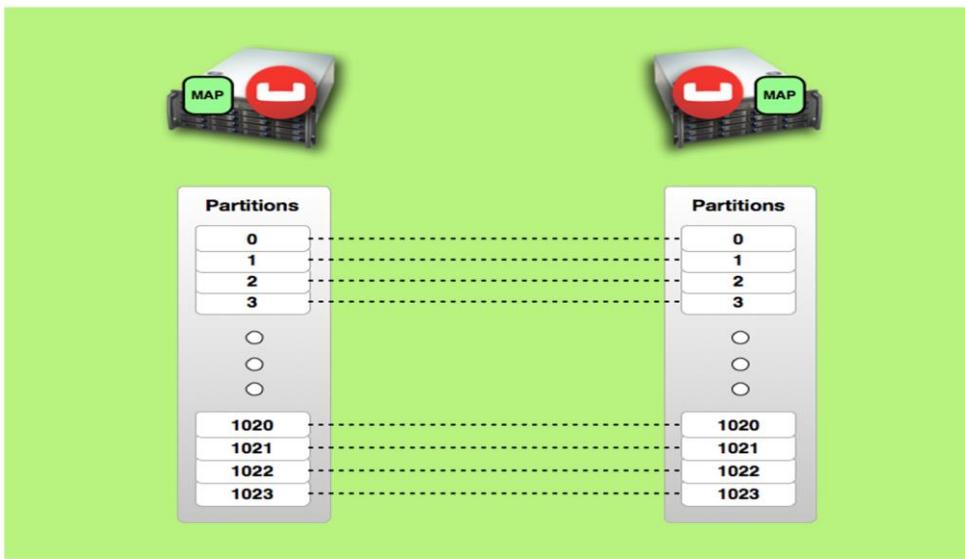
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# XDCR Ports-Connections



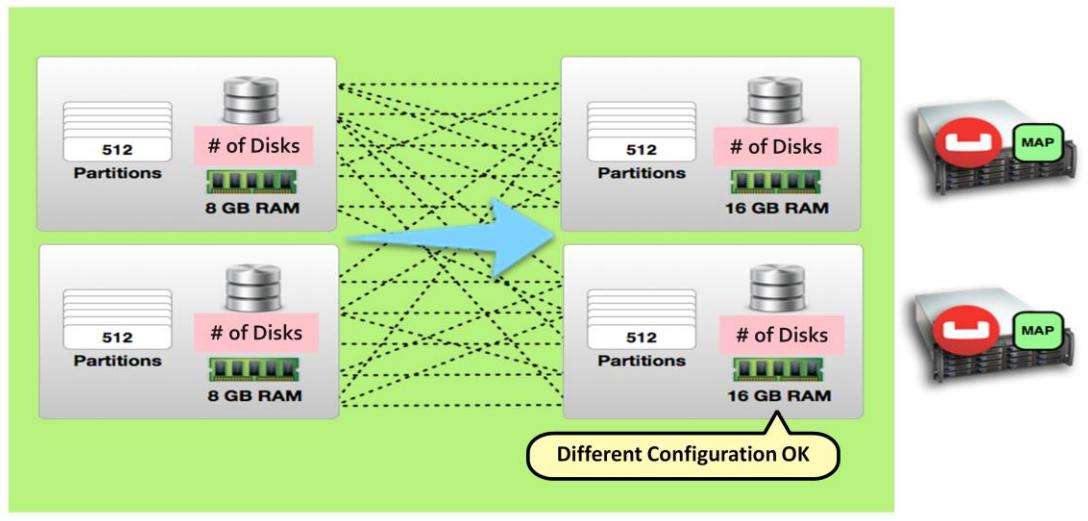
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Active-to-Active



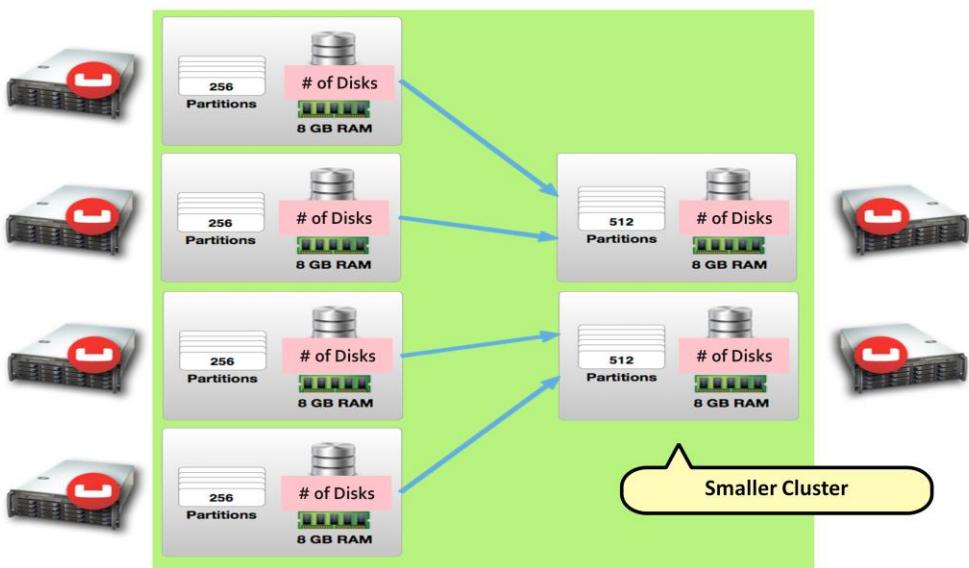
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Replication Follows Cluster Map



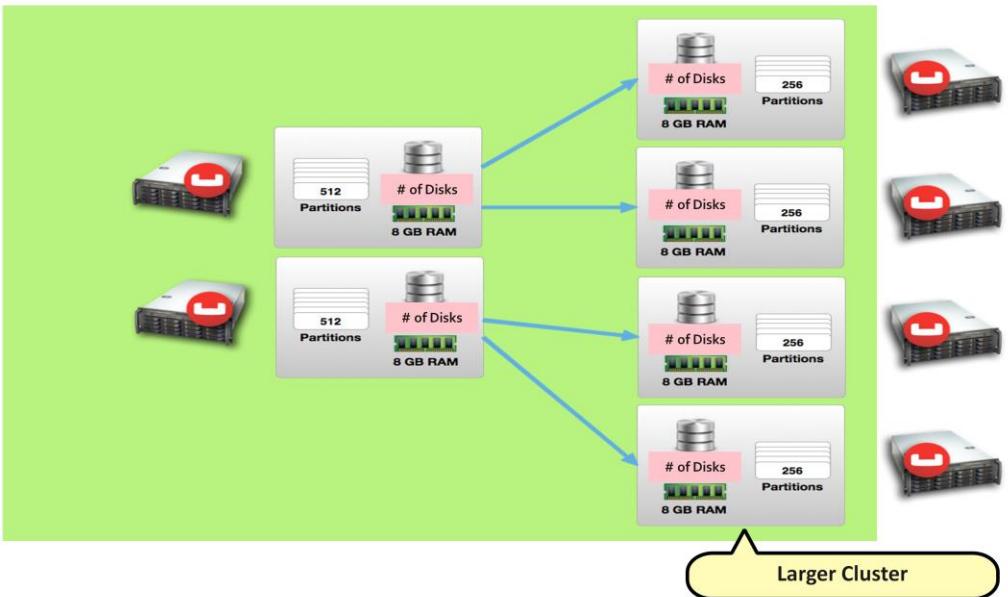
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Cluster Sizes Don't Have to Match



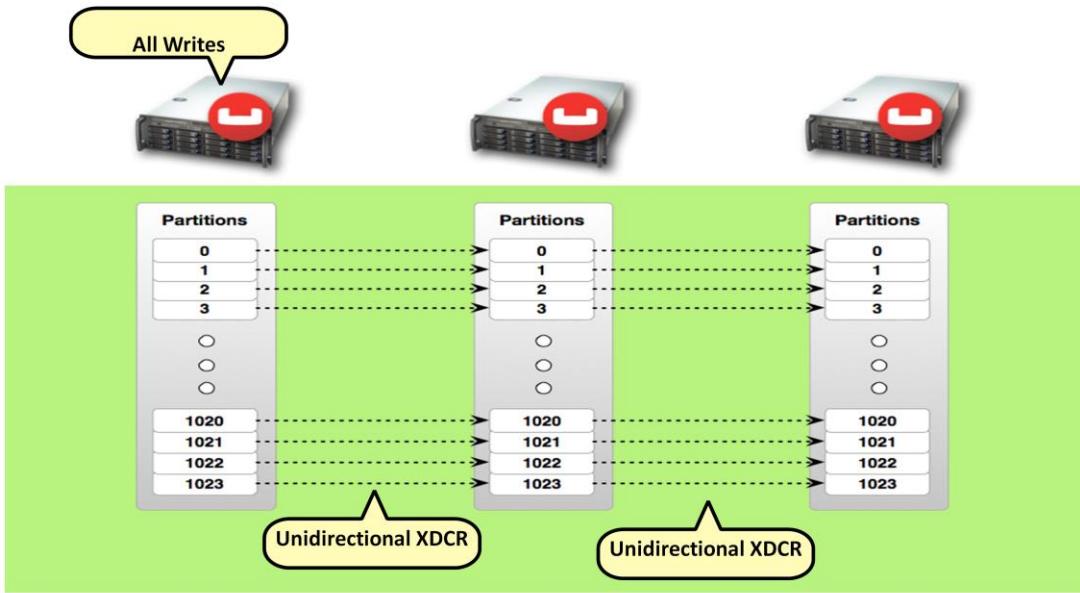
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Cluster Sizes Don't Have to Match



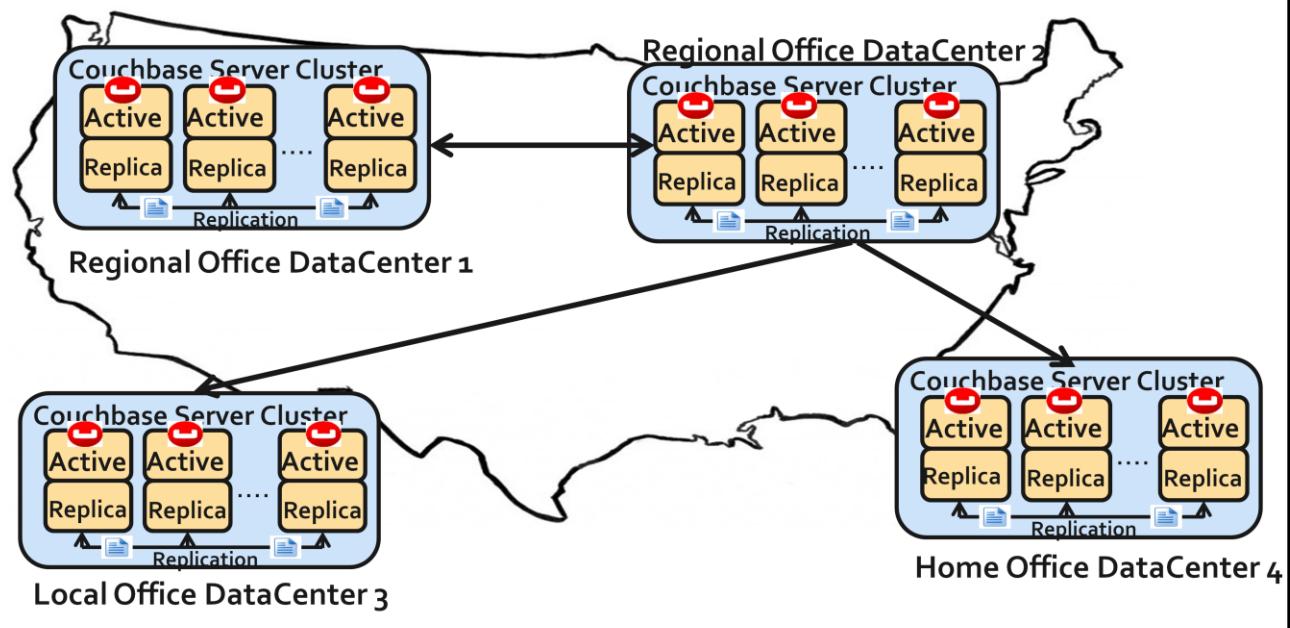
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Chaining XDCR



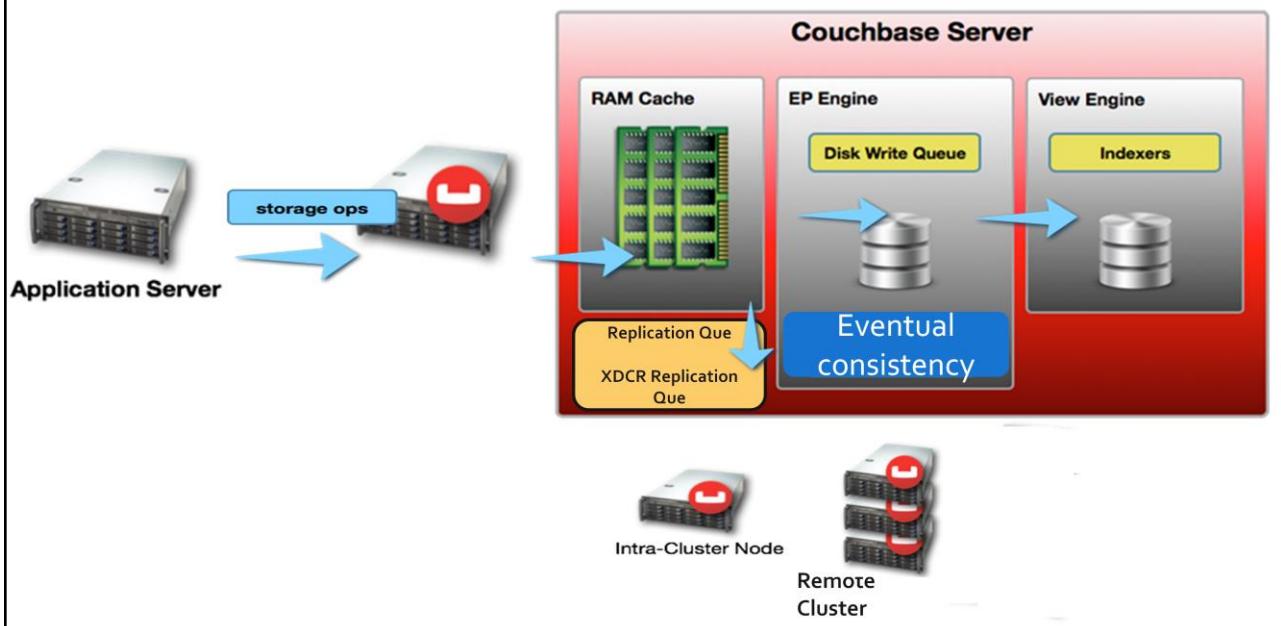
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Chaining XCDR



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Storage to XDCR Replication

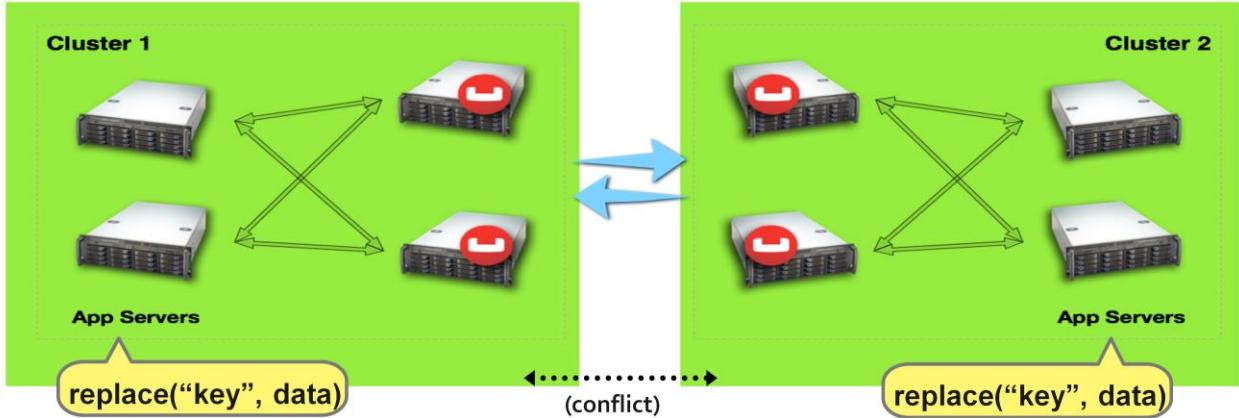


Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Bi-Directional Conflicts



What happens when you write to the same key in two clusters?



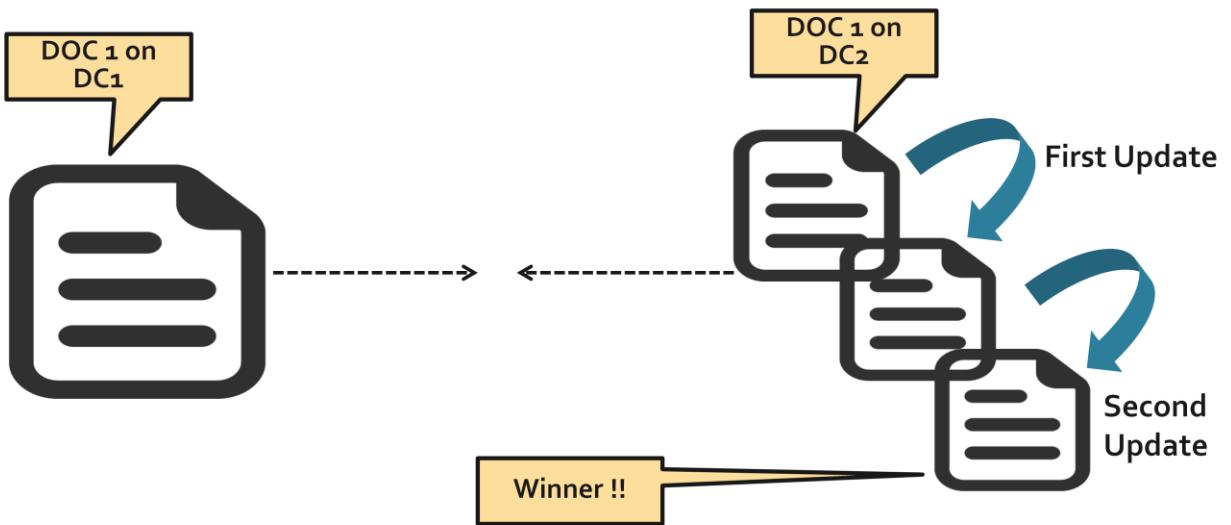
Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).



## CONFLICT RESOLUTION

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Conflict detection strategy in Couchbase Server



105

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

# Conflict Resolution Rules



- Completely Automatic, Currently Not Customizable
  - Rule 1: Revision #/Count
    - Whichever cluster has higher revision number will replicate to other cluster
    - If document is mutated in cluster with lower revision number, the document will not be put in replication queue
  - Rule 2: Metadata
    - If Revision Count equal, use CAS, Flags, and TTL
    - Basically Random
- It is not Time/NTP based



Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## Bi-Directional Caution!



**Avoid (as in don't) update the same document in multiple clusters with bi-directional XDCR, as it may not have expected behavior due to the conflict resolution rules.**

## Recommended

**Separate Key spaces (use DC prefix) so that you avoid conflicts on individual documents.**

**DC1::user:a9838-s92-s00**

**DC2::user:293ba-293-922**

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

## Flushes are not replicated via XDCR!



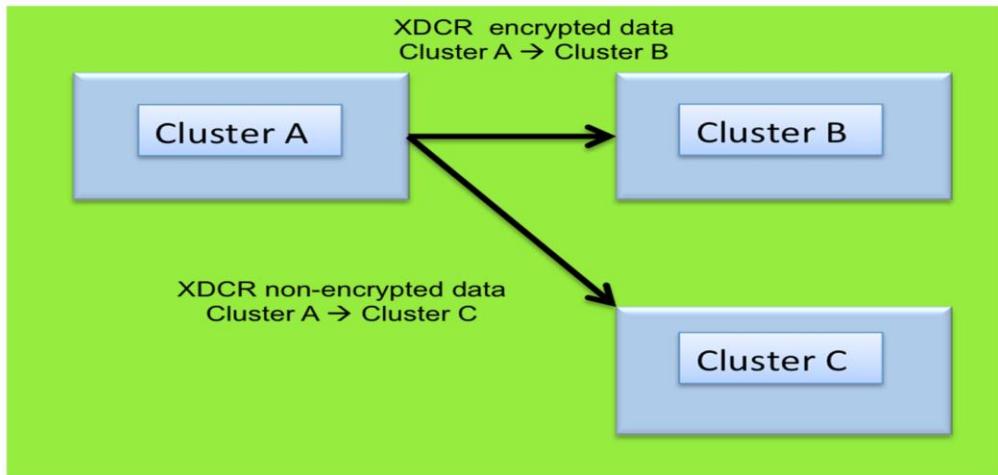
108

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Flush requests to delete the entire contents of bucket are not replicated to the remote cluster. Performing a flush operation will only delete data on the local cluster. Flush is disabled if there is an active outbound replica stream configured.

When replicating to or from a bucket, do not flush that bucket on the source or destination cluster. Flushing causes the vBucket state becomes temporarily inaccessible and results in a "not\_found" error. The error suspends replication.

# Encryption



**Both data encrypted and non-encrypted replication can not occur between the same XDCR source and destination cluster.**

109

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).

Data encryption is established between the source and destination clusters. Since data encryption is established at the cluster level, all buckets that are selected for replicated on the destination cluster are data encrypted. For buckets that need to be replicated without data encryption, establish a second XDCR destination cluster without XDCR data encryption enabled.

# Lab #6 (XDCR)



Time: 1 hour

- Create a remote 2-node Couchbase cluster
- Configure unencrypted, unidirectional replication between the 4-node and 2-node cluster
- Learn that design docs and views do not get replicated
- Examine about XDCR conflict resolution and revision counts
- Examine tombstones in the data files
- Learn about XDCR version 1 and version 2 protocols
- Advanced XDCR settings
- View outbound and inbound replication statistics
- View internal XDCR settings via REST API
- Configure encrypted, bidirectional replication between 2 clusters
- Adjust optimistic threshold setting for replication streams
- Learn how to delete replication streams

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).



# Couchbase

## Learning Services

PARTICIPANT USE ONLY IN CLASS Q3 2015

Copyright © 2014 Couchbase, Inc. All rights Reserved. NOT FOR DISTRIBUTION

Couchbase course materials are exclusively for use by a single participant in an authorized hands-on training course delivered by Couchbase, Inc. or by a Couchbase Authorized Training Partner, as listed at [www.couchbase.com](http://www.couchbase.com). Use or distribution other than to a participant in such training event is prohibited. If you believe these course materials have been reproduced or distributed in print or electronic, without permission of Couchbase, Inc. please email [training@couchbase.com](mailto:training@couchbase.com).